EDITED BY

A. A. LUMSDAINE
PROGRAM DIRECTOR
AMERICAN INSTITUTE
FOR RESEARCH
PROFESSOR OF EDUCATION
UNIVERSITY OF CALIFORNIA
LOS ANGELES

AND

ROBERT GLASER
PROFESSOR OF PSYCHOLOGY
UNIVERSITY OF PITTSBURGH
RESEARCH ADVISOR
AMERICAN INSTITUTE
FOR RESEARCH

TEACHING MACHINES AND PROGRAMMED LEARNING

a source book

DEPARTMENT OF AUDIO-VISUAL INSTRUCTION
NATIONAL EDUCATION ASSOCIATION

Copyright 1960

NATIONAL EDUCATION ASSOCIATION
OF THE UNITED STATES

Library of Congress Catalog Card No. 60-15721

Single copy, $7.50. Discounts: 10 percent on 2-9 copies and 20 percent on 10 or more copies. Address communications and make checks payable to the National Education Association, 1201 Sixteenth Street, N.W., Washington, D. C.

First printing, October 1960
Second printing, February 1961
Third printing, May 1961

FOREWORD

In April of 1959, A. A. Lumsdaine and I met for an informal talk in Pittsburgh. Neither of us, I am sure, had any idea that, at the end of our session, we would have launched the publication of a major book such as this one. During the course of our conversation I inquired as to the latest information on teaching machines—particularly how to obtain certain papers I was interested in reading. This interest, by the way, had been stimulated two years before by Dr. Lumsdaine's insistence on the importance of the teaching machine movement in the general development of what we had been calling "educational automation." Up to that point, sometime in 1957, the emphasis had rested entirely on the devices of mass instruction—television, mass-presented films, etc.

It had been difficult to get at the widely scattered literature on self-instructional devices. As is indicated in the introduction to the first section, this literature rested partly in Armed Forces reports, partly in unpublished papers which were duplicated and passed around from hand to hand. In answer to my query (as to where I could obtain some of these papers) Dr. Lumsdaine indicated that he and Robert Glaser had gathered a number of them together and had been thinking about the possibility of producing "some sort of a book of readings on teaching machines."

At that point in the conversation, the idea that resulted in this book was born and, insofar as I could speak as President-Elect, the Department of Audio-Visual Instruction of the NEA had added a new book to its publication list—a book of readings on teaching machines and programmed learning. The message was carried to Anna Hyer, the Executive Secretary of DAVI and Director of the Division of Audio-Visual Instructional Services of the NEA. She efficiently put the project in motion and now, a little over a year later, you have the result before you.

It is entirely fitting that DAVI should publish this book. Our members are technologists of the teaching profession. In his work, the practicing audio-visual director now worries about many kinds of "teaching machines"—motion picture projectors, cameras, television, tape recorders, etc. He also worries about "program" materials—films, filmstrips, tapes, TV programs. It is highly probable that, as the new teaching machines and their accompanying programs move out of the laboratory into everyday instructional use, many recurring problems of utilization, teacher education, scheduling, and even service will be dropped into the audio-visual director's lap.

Further, at the college and university level, audio-visual specialists must study and solve problems dealing with teaching machines and programmed learning. This may be due to the fact that audio-visual specialists at these levels are responsible for a wide variety of activities. Some are administrators operating complex service centers and production programs; others are engaged in teacher education and the training of other audio-visual specialists; others may do some of both and, in addition, undertake research. Within these patterns, teaching machines have already appeared. Furthermore, the boundaries of the A-V specialization are varied and shifting and interlock with several disciplines. It is no accident that the links between the audiovisualist and the experimental psychologist have been strengthened through a short but vigorous tradition. Lashley and Watson inaugurated experiments on film research; throughout the years this connection has been maintained. Hoban, for example, took his degree in educational psychology and has carried this orientation throughout his many contributions to audio-visual research. Much of the research on audio-visual media has been done by psychologists—Beck, Carpenter, Freeman, Gibson, and Lumsdaine are only a few names on the list. As an examination of the papers contained in this book will show, the development of self-instructional devices has primarily been in the hands of experimental and applied psychologists. Somewhere in the process, if history is a guide, some of these psychologists will tend to become audio-visual media specialists and some audio-visual specialists will become applied psychologists. Thus, we find Douglas Porter making an analysis of audio-visual devices in relation to a discussion of teaching machines.

Psychology does not represent the only ambiguous boundary of the audio-visual education field. Another tradition—just as strong as the psychological—binds us to the applied areas of curriculum and instruction. Again, this is easy enough to establish: W. W. Charters, one of the great curriculum leaders of a generation ago, left an indelible mark on the audio-visual field—this is the tradition of Edgar Dale and Max Corey. And curriculum and instruction, as fields within the general discipline of education, are intimately related to the problems of programing—the core of the teaching machines development.

The teaching machines described by the papers in this book are but a beginning, and the programs presented and discussed will, one day, be considered crude and unsophisticated. Some audio-visual specialists, misunderstanding, perhaps, the role of their own specialty in education, publicly worry because most of the programming so far studied has been verbal instead of pictorial. However, some of the work too recent to have found its way to publication in these pages does center around pictorial or "audio-visual" programming—for example, Keislar's use of a slide and tape machine. The issue is not verbal vs. pictorial; rather, the issue is simply that we are finally developing an instructional technology and, in that technology, self-instructional devices are going to play an increasingly important part.

If the audio-visual movement has done nothing else for American education in the fifty years of its modern development, it has attempted to improve instruction by creating, testing, and selling a gradually improving instructional technology. The days of selling have, to a large degree, disappeared. In the meantime, the technology is growing faster and faster. The audio-visual professional, as a technologist of the teaching profession, must relate to fields like psychology exactly as the medical doctor relates to his basic sciences. This interaction, I might add, is just as important for the basic scientist as for the practitioner. This is why there is much nurture for psychologists and educators alike in TEACHING MACHINES AND PROGRAMMED LEARNING.

The Department of Audio-Visual Instruction, in publishing this book, is attempting to discharge its duty to its membership, to the profession of education as a whole, and to the basic disciplines that support the profession. We have been able to do this only through the unstinting efforts of Arthur A. Lumsdaine and Robert Glaser and the support given them by the Department in the persons of Anna Hyer and Robert Snider. We are grateful, and I know that the profession will be, too.

James D. Finn, Professor of
Education, University of Southern
California, and
President, Department of Audio-Visual
Instruction

Whittier, California
June 1960

PREFACE

This volume represents the work of many contributors, to whom we should like to express our appreciation for kindly granting permission to reproduce their papers and, especially, for the tolerance they showed in permitting editorial revisions, annotations, and condensations. Special thanks are due to Sidney L. Pressey who, in addition to preparing one paper expressly for inclusion in the volume, gave generously of his time in editing a number of his previous papers and in preparing abstracts of some of the studies of his students. Our special appreciation is also due to the several other authors who wrote papers specifically for this volume or did substantial rewriting of their papers at our request.

The major task we faced as editors was that of selecting, organizing, and editing the material to provide both a comprehensive source document and a reasonably representative picture of past and current developments. The orientation and pattern of organization which we have utilized in preparing the book was influenced by numerous discussions with colleagues at the American Institute for Research and the University of Pittsburgh with whom we were working concurrently on related research projects. These include David Angell, L. J. Briggs, James L. Evans, George L. Gropper, Robert B. Hessert, Lloyd E. Homme, David J. Klaus, Harris H. Shettel, and Charles J. Stelter, to whom we are indebted for ideas which helped guide our work. We should particularly like to express our gratitude to Lloyd Homme, whose early enthusiasm and continued work in programmed learning did much to motivate our interest in this field, and to David Klaus, who has provided a critical sounding board for many of the notions which we discussed in relation to the preparation of the volume. Many of these notions, of course, stem directly from the writings of B. F. Skinner and from the earlier work of S. L. Pressey. The book was put together and the editing was done at the American Institute for Research, and we wish to express our appreciation to the Institute and, in particular, to Dr. John C. Flanagan for his encouragement of our efforts on this project.

During the period when the book was being prepared, both of the Editors were actively engaged on several research projects concerned with teaching-machine programs and with methods of training and education, under sponsorship of the United States Office of Education, the Air Force Office of Scientific Research, the Air Force Cambridge Research Center, and the Office of Naval Research. Many of the ideas and concepts which influenced the organization and editing of the volume grew out of this concurrent work. In a less direct way our thinking has also been influenced by related work in which we previously participated over a period of years under the sponsorship of various agencies, particularly the projects sponsored by the Office of Naval Research and the Air Force Personnel and Training Research Center.

A number of the original studies reported by contributors to this volume also received direct support from the Office of Naval Research, agencies of the Air Research and Development Command of the U. S. Air Force, the U. S. Office of Education, and a number of other agencies including HumRRO (Department of the Army), the Ford Foundation and the Fund for the Advancement of Education, and several industrial organizations. Acknowledgement should also be made of the support and encouragement provided on several of these projects by the University of Pittsburgh, Harvard University, and a number of other academic institutions.

Appreciation is also expressed for permission given by publishers to reprint previously published papers in this book. These publishers include the American Psychological Association, (*Contemporary Psychology*, *Journal of Educational Psychology*, and *Journal of Experimental Psychology*), *School and Society*, *Transactions of the Kansas Academy of Science*, *The Journal of Experimental Education*, *Journal of Psychology*, the Ohio State University Press, the University of Pittsburgh Press, *Harvard Educational Review*, *Science*, *Harvard Graduate School of Education Association Bulletin*, Educational Testing Service, *Educational and Psychological Measurement*, *Psychological Reports*, *IRE Transactions on Education*, *Education and Commerce*, *Engineering and Science Monthly*, *Journal of Engineering Education*, *Audio-Visual Communication Review*, the System Development Corporation, and the Wright Air Development Division of the Air Research and Development Command.

We particularly wish to acknowledge the extensive use made in Appendix I of bibliographic annotations prepared by Edward

B. Fry, Glenn L. Bryan, and Joseph W. Rigney in "Teaching Machines: An Annotated Bibliography" under the sponsorship of the Office of Naval Research and subsequently published in the *Audio-Visual Communication Review*. We are indebted to A. W. Bendig for providing technical advice on a portion of the manuscript. Thanks for provision of photographic material is extended to S. L. Pressey, B. F. Skinner, Gil Cortlett, Duette Photographers, the Department of Photography of Ohio State University, the Rheem-Califone Corporation, Western Design (U. S. Industries, Inc.), Bell Telephone Laboratories, Inc., and other organizations.

For their help in the preparation of the manuscript and the checking of references and proof, we should like to thank all of the members of the staff of the American Institute for Research and the Department of Psychology of the University of Pittsburgh who assisted in these matters. We particularly wish to indicate our thanks for the extensive work done on the manuscript by Caroline Marchese, Mary Lou Mentzer, and Kaye Hamilton. Louis Weisz was especially helpful in checking the final page proofs.

This volume could never have been produced without the very generous support of the Department of Audio-Visual Instruction of the National Education Association. The editors wish to express their warm appreciation for the support and encouragement given by Anna L. Hyer, Executive Secretary of DAVI, Robert C. Snider, DAVI Assistant Executive Secretary, and James D. Finn, DAVI President. Dr. Snider's constant cooperation and capable handling of numerous administrative details was an essential factor in the completion of the work. Special thanks are due to Gertrude Warner of the NEA Publications Division for her work in handling the final phases of checking the manuscript and readying the proofs for the printer, and to Beth Nelson who, over the entire period during which the book was being produced, worked closely with the Editors for long hours in coordinating the many details incident to the publication and supervised the preparation of proofs and bibliographic material, as well as lending a helping hand in many other respects.

A. A. Lumsdaine
Robert Glaser

Pittsburgh
July 1960

Contents

Foreword iii
Preface vi

PART I PURPOSE AND SCOPE OF THIS BOOK 1

Teaching Machines: An Introductory Overview
—A. A. Lumsdaine 5

Christmas Past, Present, and Future: A Review and Preview
—Robert Glaser 23

PART II PRESSEY'S SELF-INSTRUCTIONAL TEST-SCORING DEVICES 32

A Simple Apparatus Which Gives Tests and Scores—and Teaches
—S. L. Pressey 35

A Machine for Automatic Teaching of Drill Material
—S. L. Pressey 42

A Third and Fourth Contribution Toward the Coming "Industrial Revolution" in Education
—S. L. Pressey 47

The Value of Guidance in Reading for Information
—J. C. Peterson 52

Results of Use of Machines for Testing and for Drill upon Learning in Educational Psychology
—James Kenneth Little 59

A New Self-Scoring Test Device for Improving Instruction
—George W. Angell and Maurice E. Troyer 66

Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-Instruction
—S. L. Pressey 69

Certain Special Factors Involved in the Law of Effect
—Avery Livingston Stephens 89

PART III SKINNER'S TEACHING MACHINES AND PROGRAMMING CONCEPTS 94

The Science of Learning and the Art of Teaching
—B. F. Skinner 99

A Critical Review of a Portion of the Literature on Teaching Devices
—Douglas Porter 114

The Rationale of Teaching by Skinner's Machines
—Lloyd E. Homme ... 133

Teaching Machines
—B. F. Skinner ... 137

The Use of Teaching Machines in College Instruction
—B. F. Skinner and James G. Holland ... 159

An Application of Recent Developments in Psychology to the Teaching of German
—Charles B. Ferster and Stanley M. Sapon ... 173

A Report on Instructional Devices in Foreign Language Teaching
—Douglas Porter ... 186

Teaching Machines
—Douglas Porter ... 206

Teaching Machines: An Application of Principles from the Laboratory
—James G. Holland ... 215

Report on the Initial Test of a Junior High-School Vocabulary Program
—Susan R. Meyer ... 229

An Exploratory Investigation of an Apparatus for Studying and Teaching Visual Discrimination, Using Preschool Children
—Wells Hively ... 247

PART IV CONTRIBUTIONS FROM MILITARY AND OTHER SOURCES ... 257

Teaching and Educational Inventions
—Ibert Mellan ... 265

The Tab Item: A Technique for The Measurement of Proficiency in Diagnostic Problem-Solving Tasks
—Robert Glaser, Dora E. Damrin, and Floyd M. Gardner ... 275

Automatic Tutoring by Intrinsic Programming
—Norman A. Crowder ... 286

Two Self-Instructional Devices
—Leslie J. Briggs ... 299

Heuristic Discussions of Psychological Bases for the Conduct of Training by Automatic Devices: A Functional Diagram and an Approach to Research
—Ernst Z. Rothkopf ... 305

A Proposed "Automatic" Teaching Device
—Paul K. Weimer ... 311

Some Research Problems in the Design of Materials and Devices for Automated Teaching
—Ernst Z. Rothkopf ... 318

Problems in Simulation and Programming in the Design of Complex Trainers
—Leslie J. Briggs ... 329

Electronic Keyboard Teaching Machines
—Gordon Pask ... 336

Adaptive Teaching with Adaptive Machines
—Gordon Pask ... 349

A New Technique of Education
—Simon Ramo ... 367

Technology and the Instructional Process
—James D. Finn ... 382

PART V SOME RECENT WORK ... 395

Teaching Machines and Human Beings
—John W. Blyth ... 401

The Earlham College Student Self-Instructional Project: A First Quarterly Report
—John A. Barlow ... 416

Teaching Machines
—D. J. Mayhew and A. F. Johnson ... 422

The Development of Understanding in Arithmetic by a Teaching Machine
—Evan R. Keislar ... 425

An Evaluation of Textbooks in Terms of Learning Principles
—Robert Glaser, Lloyd E. Homme, and J. L. Evans 437

A Preliminary Investigation of Variation in the Properties of Verbal Learning Sequences of the "Teaching Machine" Type
—J. L. Evans, Robert Glaser, and Lloyd E. Homme 446

Results of an Initial Experiment in Automated Teaching
—J. E. Coulson and H. F. Silberman ... 452

A Study of Teaching Machine Response Modes
—Edward B. Fry ... 469

Contents

On the Relevance of Laboratory Investigation of Learning to Self-Instructional Programming
—Thomas F. Gilbert 475
Problems in Programming Verbal Learning Sequences
—Lloyd E. Homme and Robert Glaser 486
Some Perspectives and Major Problems Regarding "Teaching Machines"
—Sidney L. Pressey 497
Error Responses and Reinforcement Schedules in Self-Instructional Devices
—Abram Amsel 506
Some Issues Concerning Devices and Programs for Automated Learning
—A. A. Lumsdaine 517
A Functional Analysis of Self-Instructional Devices
—W. J. Carr 540

CONCLUDING REMARKS 563

APPENDICES 573

Introduction 574
Appendix I 581
Appendix II 691
Index 721

PART I

Purpose and Scope of This Book

THE PURPOSE of this book is to provide a comprehensive reference source on teaching machines and the techniques of instruction that are associated with them. Teaching machines and related self-instructional devices have attracted increasingly wide attention in the past few years as new media of instruction. Much of this attention has centered on the nature of specific devices for presenting instructional materials to the individual learner; however, the organization or programming of the subject matter these devices are designed to present has received somewhat less public attention. This volume deals with both of these highly interrelated aspects of an evolving educational technology, and its title—TEACHING MACHINES AND PROGRAMMED LEARNING—is intended to reflect this dual emphasis.

Although the first devices that we would today call "teaching machines" appeared over 30 years ago, they attracted relatively little attention for more than two decades. Recently, however, interest has suddenly become widespread, and developments have been occurring at a greatly accelerated rate. This upsurge of interest has created a demand for a convenient and comprehensive source of information on what has been happening and what has been written in this rapidly expanding field.

Educators and psychologists, in increasing numbers, have been seeking information about activities in this field. This effort has encountered considerable difficulty because of the relative inaccessability of much of the material. Published articles in this field have appeared in a widely scattered periodical literature. A number of the more recent contributions have not been published. Dissemination of information has depended largely on word of mouth and personal correspondence. This has been a quite inefficient process, involving much overlapping of effort, and has also imposed a heavy burden upon the authors of papers.

Available supplies of reprints or copies of unpublished papers have frequently been exhausted.

Many individuals have independently compiled bibliographies and private collections of published and unpublished papers. Several annotated bibliographies have recently been given limited distribution in mimeographed form (e.g., Darby, 1959; Fry, Bryan, and Rigney, 1960). The only published book to appear on this subject thus far is limited to papers and abstracts prepared for a symposium at the University of Pennsylvania sponsored by the Air Force Office of Scientific Research (E. H. Galanter, editor. *Automatic Teaching: The State of the Art.* New York: John Wiley & Sons, 1959). Of necessity, this collection does not provide access to original background sources that would afford broad coverage and perspective.

The present volume does not undertake to present a systematic treatise on teaching machines and programmed learning. In the opinion of the editors, such an effort, in all likelihood, would be premature at this time. Rather, this book is what its subtitle indicates: a *source book* of original papers that the student of the new field will want to consult. The major aim of this publication is simply to bring together these widely scattered reference sources and to make them conveniently available in a single volume. However, an attempt has been made to structure the field as much as possible by organizing the papers collected here in terms of the main lines of development leading up to the present state of the art.

The editors have sought to include most of the basic publications of leaders in the development of teaching machines and self-instructional learning programs. These articles span the field from the first pioneering paper published by Sidney L. Pressey in 1926 up to interim reports on present ongoing projects. Those papers that represent major landmarks of historical or systematic importance have, in general, been reprinted without editorial condensation. Some of the other reports have been edited in varying degrees to eliminate redundant portions or to delete peripheral material. One systematic exclusion has been the omission of those papers which are included in the recent collection of symposium papers edited by Galanter, since these are now readily available. However, abstracts of these papers have been included in Appendix I. In choosing among papers of the many recent contributors to the field, the editors have been faced with a difficult task of selection. A number of individuals have recently prepared excellent papers summarizing past developments and interpreting or discussing these. Limitations of space alone have prohibited the inclusion of all these recent summaries. Moreover, the newness of the field understandably has resulted in a high degree of overlap among these papers in their summarization of past work. However, available papers which for these and other reasons have not been included in the text of the volume have been abstracted in Appendix I.

In Appendix I the editors have endeavored to identify all published papers known to them that deal directly with teaching machines and programmed learning. Also abstracted here are a number of papers that have been presented at symposia or that have been reproduced as reports or memoranda by government agencies and other institutions, without general publication. The editors have felt that the newness of the field and the rapidity with which it is developing make it useful to abstract such informally circulated papers, some of which have been cited in published articles.

This volume has been organized in five parts. In Part I the editors have attempted to provide a preliminary overview and survey of developments by including two recent reviews. The first of these, adapted from a recent article by Lumsdaine, identifies basic characteristics common to all devices currently classed as teaching machines. It illustrates these characteristics by describing a number of devices that have been developed and also provides a brief characterization of some of the techniques of programming materials for use in various devices. The other introductory paper reprints a recent review by Glaser, in which he has viewed current and past developments in historical perspective. In this review, the necessity for improvement in educational technology has been related to the promise that the evolution of the teaching machine may hold for meeting this need. This has been done by a review of several key articles. These represent, first, the pioneering contributions of Sidney L. Pressey; second, the recent leadership and major contributions to the field provided by B. F. Skinner; and, third, some speculations about the future of education in an increasingly technological society as projected by Simon Ramo.

The original papers, representing the basic source material of this book, have been grouped in four parts which reflect the major lines of development in the field. Part II begins with the early work initiated by Pressey 35 years ago, which for some years went generally unrecognized. This section reports work along the lines initiated by Pressey, continuing up to the mid-fifties. Part III presents the basic articles of Skinner, who in 1954 pointed out that his previous work on the analysis of behavior was applicable to the development of teaching machines. This section includes work that has continued under his general guidance up until the present time. Part IV contains contributions, from a variety of sources, which are related to the theme of this volume but which are not directly continuous with the work of Pressey and Skinner. A number of the articles in Part IV grew out of the work of experimental psychologists on problems of training and instruction in the military services. Part V represents a sample of recent activities and current trends of thinking on teaching machines and programmed learning, which have been largely stimulated by the work of Skinner and his associates.

For each of these major parts of the book, the editors have provided a brief introduction that describes some general characteristics of each group of papers, notes significant features of individual papers, and seeks to provide a useful perspective. Throughout the text, references are provided to related papers which have not been reprinted in full but which have been abstracted in Appendix I. (In Appendix I no internal structuring in terms of chronology or line of development has been used; to facilitate search, abstracts of papers or annotated titles are presented in straight alphabetical order by author and year.)

One systematic editorial change has been made in all the papers for the convenience of the reader; namely, references to published literature contained in footnotes or bibliographic lists in the original articles have been reduced to a common style of reference, and the frequently overlapping bibliographies in the original papers have been collected in one comprehensive reference list at the end of the volume. Reference citations in the text are made by author and date, so that these can be readily located in the single reference list (Appendix II).

Teaching Machines: An Introductory Overview

A. A. LUMSDAINE

THIS PAPER discusses some of the developments that have occurred during the past few years in the construction of mechanical devices for individual self-instruction. It has become increasingly apparent that such devices and the instructional programs they present are likely to have very important implications for educational practice. It is even possible that, in the relatively near future, they could modify profoundly many of our basic conceptions of instructional method and of the role of teaching aids and audio-visual materials in the educational process. . . .

Teaching machines and the instructional programs used in them have developed partly on the basis of a large body of work by psychologists in conducting experimental studies of learning, and the features the machines incorporate reflect important principles of effective learning to which these studies have pointed. The present paper does not attempt to summarize this underlying experimental data but instead concentrates on developments in devices and materials. In these developments, an attempt has been made to implement principles of effective learning in ways which will be indicated in relation to features of the various devices.

Despite great variation in complexity and special features, all of the devices that are currently called "teaching machines" represent some form of variation on what can be called the tutorial or Socratic method of teaching. That is, they present the individual student with programs of questions and answers, problems to be solved, or exercises to be performed. In addition, however, they always provide some type of automatic feedback or correction to the student so that he is immediately in-

---

With the exception of minor editorial changes and the addition of one short section, this paper is reprinted from an article that appeared in the Summer 1959 issue of the *Audio-Visual Communication Review* under the title "Teaching Machines and Self-Instructional Materials."

formed of his progress at each step and given a basis for correcting his errors. They thus differ from films, TV, and most other audio-visual media as ordinarily utilized, because of three important properties.

First, continuous active student response is required, providing explicit practice and testing of each step of what is to be learned.

Second, a basis is provided for informing the student with minimal delay whether each response he makes is correct, leading him directly or indirectly to correction of his errors.

Third, the student proceeds on an individual basis at his own rate—faster students romping through an instructional sequence very rapidly, slower students being tutored as slowly as necessary, with indefinite patience to meet their special needs.

The devices thus represent a way of providing a pre-programmed study-practice combination which simulates, in partially or fully automated fashion, the functions of a private tutor in recitation and practice, with immediate correction of errors and feedback to the student. [The brief overview of several devices, given here for general orientation, is supplemented by more detailed descriptions given in the original papers presented later in this volume.]

The device shown in Figure I illustrates many of the features of teaching machines. The panel at the left of Figure I gives a close-up view of the display with which the student, shown working at the right, is faced. This device was designed as part of an Air Force research and development program and is described in more detail by Briggs (1958). The device, called the Subject-Matter Trainer, represents a further evolution of some of the simpler devices which were developed and studied earlier by Sidney L. Pressey and some of his students at The Ohio State University. About half a dozen of these machines have been constructed and tried out experimentally for instruction of technical specialists in the Air Force.

In essence, what this machine does is to present a predetermined sequence of questions, one at a time, in the aperture at the left of the panel. The questions can be answered by selecting any of about 20 response choices, which are displayed either verbally or pictorially on the large panel. Beside each answer choice is a green indicator light and a push button by which the student may identify and record the answer he selects

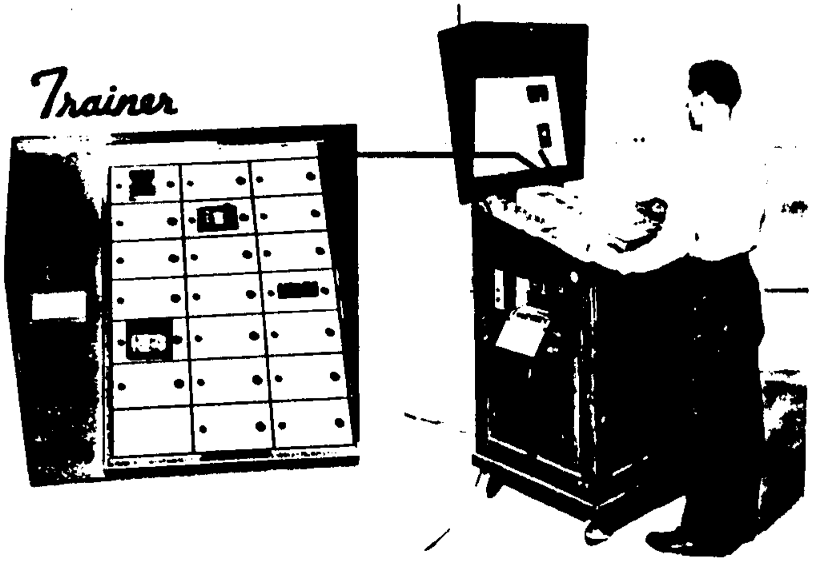


FIG. I. "Subject-Matter Trainer." This is one of a class of automatic self-instructional devices or "teaching machines" suitable for training in identifying of components, terms, and other "one-to-one" associations. The student is continuously tested as he learns.

as the correct choice. The device has considerable flexibility and can be programmed for use either as a testing machine or for any of several modes of operation for self-instruction in a variety of subject matters. It can be programmed, for example, so that the green light indicating the correct answer choice comes on at once the first time through, so that the student is guided to the correct choice. Later, he may be required to decide on a choice; if correct, the light beside the answer he has chosen so informs him, while if incorrect, the light next to the answer that *is* correct comes on. Still later, he may be required to keep trying until he gets the correct answer. All the while, his performance is recorded and his errors totaled, so that he is continuously tested while he is learning. As he progresses, the machine can drop out the items he has mastered so that the student's time and interest are not dissipated by unnecessary repetition.

This is obviously a rather expensive device. It was built, however, for Air Force training situations in which its cost would be very minor in relation to the total training cost for technicians who used it, not to mention the cost of errors that ineffective training could occasion in the operation of a $10 million bomber or missile. In addition, this particular device was designed to have considerable flexibility and a number of modes of operation because it was also planned as a research tool for investigating various features, data on which would affect the design of simpler future devices for standard training situations.

Let us now take a look at some of the smaller and somewhat simpler devices that have been developed for use in civilian educational situations.

The device at the left of Figure II might be thought of as an ancestor of all subsequent teaching machines. It was devised by S. L. Pressey (1926, 1927, 1950) over 30 years ago. Originally, it was conceived primarily as an automatic *testing* device, but it soon became apparent that its *instructional* properties were also of great interest. In the sketch at the left of Figure II is seen a window, at the left, in which a multiple-choice question appears. For example, in a course in elementary physics or electricity, the question might ask the student to identify the proper symbol to represent electric current or to solve a simple problem applying Ohm's Law. His basis for doing so might be the prior study of his textbook, or one of Harvey White's physics TV/film lessons, or some other source—including what he had learned from answering preceding questions in the series. In any case, suppose he chooses the proper answer and presses the third push button, corresponding to this answer. If he does so, the machine signifies the correctness of his answer by presenting the next question or problem. But if he is wrong the question stays in the window, an error is tallied on a counter at the back of the device, and he must try again until he selects the right answer before the next question will be presented.

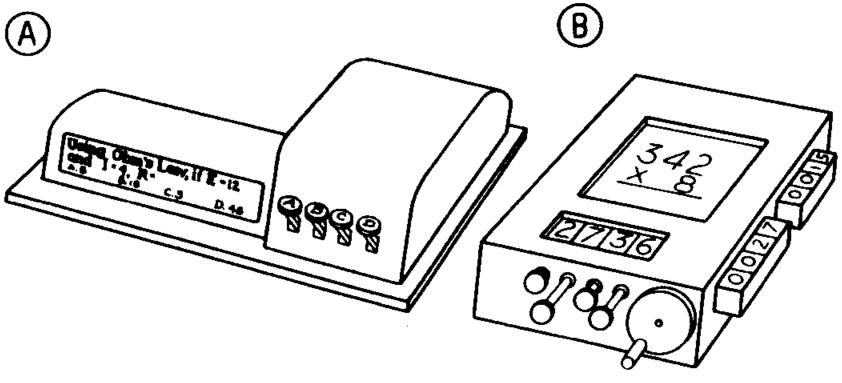


FIG. II. Two simple individually paced "teaching machines" which have automatic self-scoring features as in the device shown in Figure I.

Some of the other devices developed by Pressey and his students were further simplified so that the four-choice answer apparatus was used with an ordinary mimeographed test form. This introduces the disadvantage that there is less control over the sequencing of information presented to the student—though these devices were keyed so that the student had to answer the questions in prescribed order, he still could look ahead or back in the series of questions and thereby obtain uncontrolled or non-programmed clues to the answering of a given question. How important this disadvantage really is has not, curiously enough, been given much attention, and research is needed to investigate this problem.

In similar devices developed by Pressey and others, the mechanical gadgetry was simplified by the use of punchboards to replace the mechanical push-button arrangements. In the punchboard device (see Pressey, 1950) the student simply punched his pencil point through one of four holes . . . corresponding to the answer he selected. If right, a key-card under the punchboard holes would let the pencil . . . perforate the answer sheet; if wrong, the pencil point would be stopped, indicating an error to the student—who then tried another answer. Still a different device for self-scoring and immediate feedback was devised by Peterson (1930) by use of chemically treated answer sheets and a special marking pencil. . . . [The correct answer, when marked by the student, was identified by the marked area's color.]

For some reason these multiple-choice devices, though shown by experimental tests to be quite effective aids to learning, never caught on widely during the two decades following the initial development of Pressey's first machines. This may be partly because the times were not ripe for their acceptance and partly because they were conceived primarily as testing devices and only secondarily as teaching machines. Also, it may have been due in part to some of their inherent limitations. One of these limitations is that, as multiple-choice devices, they appeared to be limited to *recognition* responding, rather than permitting the student to compose or construct his own response. (The latter advantage is, as a matter of fact, the subject of some controversy among teaching-machine designers at the present time and is not fully resolved by the research data thus far available.)

A device that permits constructive responding is shown at B in Figure II. This particular device, designed by David Zeaman at the University of Connecticut, is similar to one of the machines earlier constructed by B. F. Skinner at Harvard. It is a special-purpose machine designed for teaching elementary arithmetic to grade-school children.

A problem appears in the large window on the face of the machine. The problem might be of any degree of difficulty from 2 x 2 up to a problem in long division or square-root derivation. The student, after whatever supplementary mental or scratch-paper figuring he needs, composes his answer by moving the four plungers at the front of the machine. These operate the dials to show his answer in the aperture below the problem. (In this case, 2736, the correct answer, has been set in.) When he is satisfied with his answer, the student turns the crank at the right. If the answer is correct, the machine advances to the next problem by moving a tape on which the problems are printed. If wrong, it scores an error and leaves the problem unchanged so that he must try again until he gets the correct answer before he can proceed.

Several similar devices have been built, some with an extended repertoire of response-construction including letters as well as numbers to permit their use for teaching spelling and other verbal subjects. In these devices, as in Pressey's four-answer multiple-choice machine and Briggs's more elaborate 20-answer multiple-choice apparatus, the correctness of the answer is scored automatically.

Most devices with more extensive response repertoires have sacrificed automaticity of scoring to keep the instrumentation from becoming too complex (or even from exceeding the present state of the art). In these devices, the student himself judges whether the answer he has given is correct and instructs the machine accordingly. Perhaps the best known device in use thus far is one devised by Professor Skinner at Harvard and shown at A in Figure III.

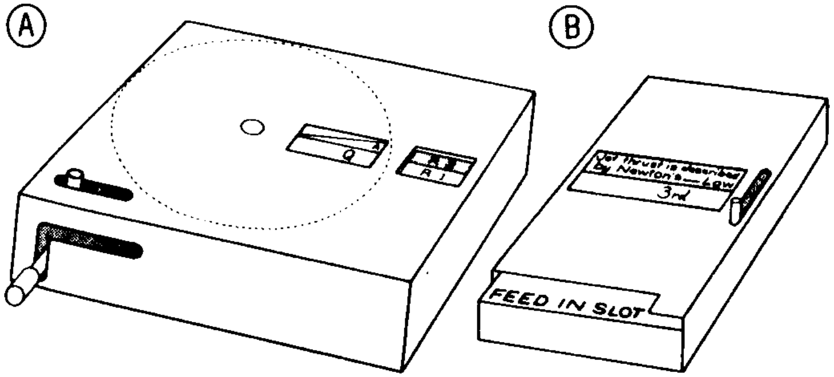


FIG. III. Two "teaching machines" using written responses which are scored by the student on the basis of comparing his written answer with the correct answer.

This machine (Skinner, 1958c) employs questions that are printed on a disk mounted under the cover of the device and shown in the sketch by the dotted circle. Only one question at a time is visible to the student as it appears in the window marked "Q." The adjacent correct answer, just above at A, is not visible until the student has completed his answer. He does this by writing it on an exposed frame of a paper tape at the right, at R1. When he has written an answer he is satisfied with, the student raises the lever at the left front of the machine. This simultaneously exposes the correct answer at A and moves the answer the student has just written up under a transparent cover to R2, where he can see it but can no longer change it. The student now compares the correct answer with his own and decides whether he was right or wrong. If he decides he was correct, he moves the lever over to the right. This marks his answer tape to show that he scored himself correct on that question and also operates a mechanism that reduces the number of times that question will appear again. (The number of correct answers required for each question may be either two or one with this machine, depending on the way it is set.) Whether the answer was right or wrong, the next question appears as soon as the lever is returned to the lower position, and the student goes on with the new question.

The most elaborate aspect of the internal mechanism in this device is the part that allows for dropping out questions that have been answered once or twice correctly. One of the numerous questions on which we need research as a basis for future designs is that of how important this drop-out feature is for various kinds of subject matter and kinds of students. This question interacts with the number of questions included in any one cycle. For the machine in Figure III-A, the cycle comprises the 30 questions on one disk. In the much simpler device, shown at B in Figure III, adjacent to the Skinner machine, a shorter cycle of only five questions is used, with no drop-out arrangement. In this device, built by Douglas Porter (1958c, 1959) at Harvard, questions on mimeographed or dittoed sheets are fed into the device at the bottom and appear in the window. The student marks his answer directly on this sheet. For example, in a program used for teaching elementary spelling in a nonrote, context approach, the device might present the sentence, "Rain came down in a steady drizzle," with the omitted letters T and D in the word "steady" to be filled in by the pupil. The student fills in the letters that he thinks belong in the blanks, then pushes the lever at the right. This advances the sheet about an inch up and exposes the correctly completed word. As in Skinner's machine, the blanks where student's response was entered have now been covered by a transparent plastic or glass cover, so that he cannot cheat by waiting until the correct answer is revealed before making his attempt. This device is not fully self-scoring but does provide important advantages in a low-cost, fairly compact device. Other devices of still greater simplicity also have been designed that still retain some of the key features of their more elaborate counterparts. One of the simplest is a so-called programmed textbook devised by Professors Glaser and Homme at the University of Pittsburgh.

Glaser, Homme, and Evans (1959) employ response-and-correction sequences provided by a special, nonconventional form of "programmed" question-and-answer book, in which the answer to each question is always given on the following page. These devices are especially attractive in terms of cost, but they lack some of the features of stimulus control, and perhaps some of the fun of learning and being quizzed by a semi-automatic "gadget" or device—which is part of the appeal that slightly more complex devices seem to have for students. Also, they do not have the self-scoring features that are capable of relieving the teacher of the drudgery of quiz-scoring; and since the answer appears on the succeeding page in each case, they are not cheat-proof even as practice devices.

How important the self-scoring and cheat-proof features are is partly an administrative question—i.e., how much do we want to pay for getting an automatically scored quiz paper—and partly a question for research—i.e., how much is learning benefited if the device is constructed so that the student cannot readily beat it? There is some evidence on the latter question, but more research will be needed before we have anything approaching a definitive answer.

The problem of making a device fully automatic and cheat-proof runs counter to the requirements for making it inexpensive and yet allowing a variety of kinds of free response to be made by the student. One experimental research device constructed by E. Z. Rothkopf (1958b) to experiment with a variety of response repertoires is shown in Figure IV.

In this device, extended multiple-choice capabilities are provided, together with automatically scored verbal-composition responses of any of the 26 letters and 10 numbers in a word or expression of up to some 30 symbols. These are entered through the use of the sliders at the right, much as in the Zeaman numerical device that was shown in Figure II-B. In addition, a self-scoring tracer stylus is provided, which can automatically score the correctness of a student's response in tracing a route on a map, following an electrical circuit diagram, constructing a simple drawing, or the like.

The Skinner and Porter machines and the Glaser and Homme so-called "paper machine," or programmed Socratic textbook, have been tried out with a number of subject matters experimentally—the former in full-year courses in some cases, the latter for shorter segments of a few experimental lessons.

As was stated earlier, all of the developments in pre-programmed self-instructional materials, and devices for presenting them, are based fundamentally on Socratic question-and-answer or problem-and-solution methods of teaching. Typically, they proceed in small steps of graded difficulty, so that mastery of

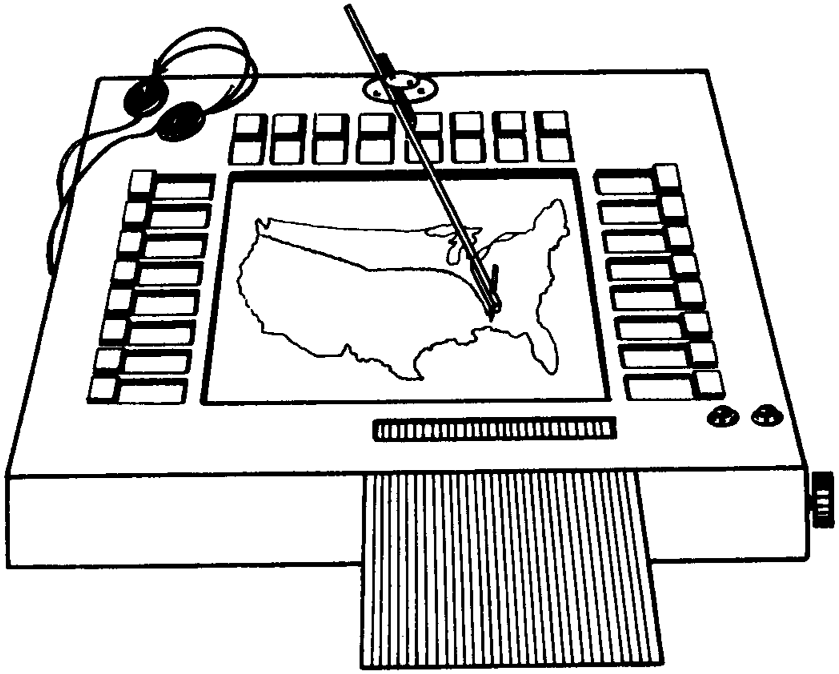


FIG. IV. An experimental teaching device designed for research study of several modes of student response (see Rothkopf, 1958).

concepts, understanding, and skills are gradually built up as the student proceeds through the program. An element that is probably crucial in their effectiveness is that of nearly continuous active student response that is monitored, guided, and corrected at each step in such a way as to insure mastery of one step before going on to the next.

In addition, optimum pacing to fit each student's rate of progress can be provided, and the sequencing of instructional materials and amount of repetitive practice employed can be governed by the proficiency shown at each step by the individual student. The slower learner is thus gradually brought up to a satisfactory level of accomplishment, while the quicker student, after reaching a satisfactory level of attainment as fast as his abilities permit, can then, if desired, proceed to more advanced material.

The programs of self-quizzing, self-teaching items used in practice machines and related devices are of great interest, and the design of these programs is the most crucial factor in the successful utilization of automated individual-tutoring methods. Perhaps the most important single factor in constructing programs is the use of prompts or coaching cues, of varying degrees of subtlety and indirectness, which are used to keep the learner responding correctly and to keep up his motivation by insuring successful performance. The programming of such cues or prompts is arranged with initial maximum use to insure elicitation of the desired responses by the students, followed by so-called "fading" or "vanishing" of the prompts in later items of a sequence so that the learner is gradually freed from dependence on them.

In the instructional programs devised by Skinner, Glaser, Homme, Porter, and others, great pains are taken to develop desired response patterns by using a wide variety of contextual and semantic prompts. One simple example of the kinds of verbal prompts that can be used is partial presentation of a word, with omitted letters to be filled in—the letters present varying from nearly all present (or maximal cueing) initially to all, or nearly all, absent (or minimal cueing) terminally.

For programs that are designed to build up mastery of complex concepts rather than to develop simpler skills, other techniques are useful. Two examples are presented in Table 1.

Table 1. SOME PROMPTING TECHNIQUES USED WITHIN FRAMES OF TEACHING MACHINE PROGRAMS

I. EXAMPLES OF PROMPTING BY ANALOGIES OR EXAMPLES:
Just as a rise of temperature increases gas pressure, a rise in filament temperature will.........?........electron emission from the filament of a vacuum tube.

II. EXAMPLE OF PROMPTING BY WORD HABITS:
When we start to trace a signal from the antenna to the viewing tube, the ....?........ component we encounter is the RF detector.

In the first example, analogy is used to get the student to emit the appropriate response in context for the first time. Almost inevitably, he will respond with the desired term to complete the statement that a "rise in filament temperature will *increase* electron emission from the filament."

In the second example, pre-existing word habits are used in a different form of cueing: The structure and meaning of the sentence make it very likely that the student will respond correctly by saying . . . "The *first* component we encounter is the RF detector."

Another general method of providing prompts to foster correct responding in context is through dependence on stimuli provided in immediately preceding frames of the sequence.

Table 2. PROMPTING BY CUES FROM PRECEDING ITEMS IN A TEACHING MACHINE PROGRAM

62. A ....?.........., used to control voltages, consists of two coils of wire, wound on a core.

63. Input and output voltages of a transformer are governed by the number of turns of wire in the primary and secondary ....?..........

64. The ratio of the number of turns in the ....?........ and in the ....?........ coils determines the ratio of input and output voltages.

Table 2 shows a brief three-item excerpt from a program in electrical fundamentals. The items would normally be seen only one at a time. Assume the student responds correctly to Item 62 with the word *"transformer"* on the basis of what he has learned in preceding sequences. If he does not, the correct answer will, of course, be provided after he has tried to answer the item. Cues for the response *"coils"* for Item 63 are provided by the statement just previously read in Item 62; similarly, the terms *"primary"* and *"secondary"* required in Item 64, when the student gets to it, will have just been used in Item 63, and so on. A number of verbal prompting techniques have been described by Smith (1959). The techniques of providing prompting cues may be subtle or more obvious and can be used for eliciting responses to non-verbal stimuli as well as for purely verbal patterns. Instructive examples are cited by Skinner (1958c) and Holland (1960).

It is clear that a great deal of ingenuity and skill may be expended in constructing good programs and also that the design of both devices and programs suggest a number of problems on which research is needed to provide firmer guidelines for future development. Some of the more important problems are the question of the size of steps in the programs, the optimal use of prompts, and the rate at which the prompts are to be eliminated or vanished.

The principles used in most of the programs constructed by Skinner, Gilbert, Homme, Glaser, Porter, and others appear very reasonable but have not yet been subjected to clear-cut experimental tests. Some highly suggestive evidence is available from studies of active student response in learning from films. For example, an Air Force study by Kimble and Wulff (1953) found that increasing the amount of coaching and guidance furnished to students in learning to use a slide rule resulted in superior learning. Other studies summarized by Allen (1957) and in a recent paper by Lumsdaine (1959a) document the value of frequent guided response by the student. In experimenting with programmed materials for teaching several subject matters, Evans, Glaser, and Homme (1959) at the University of Pittsburgh have found that, up to a point, the smaller the steps in the program, the better. How to determine that point is at present pretty much an empirical matter, and we need to develop a theory and procedures for predicting the best size of step to use with any subject matter.

Some very suggestive evidence related to the vanishing technique, in the case of procedural learning, was obtained in an Air Force-sponsored program of research through the theoretical and experimental work of Maccoby and Sheffield (1958), who concluded that transition from initially short demonstration-and-practice segments to larger segments later on would produce better terminal performance in a nonprompted test situation than consistent use of either long or short segments throughout. (In this case the larger segments provide more remote, thus less direct cueing, for the steps of the procedure.)

The effective use of short and variable length sequences for individual instruction by filmed demonstration requires projection equipment with characteristics that permit its use as a good teaching machine and that are not at present provided by commercially available projectors. One suitable device that exists at present only in a single prototype model is shown in Figure V.

This device (see Hoehn and Lumsdaine, 1958) shares with the teaching machines discussed previously the characteristics of presenting materials in a fixed sequence, at a pace that can either be preprogrammed or determined by the responses of

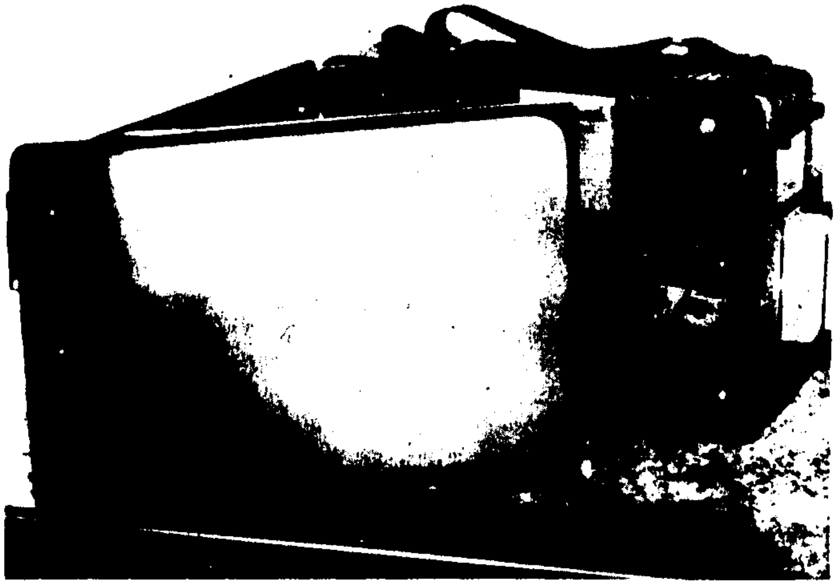


FIG. V. Magazine-loaded "automated" projector. This prototype device was designed by the Air Force to provide "automated" presentation of demonstration film sequences, with flexible access to any desired film through simple insertion of prethreaded magazines.

the individual learner using it. In this case, however, the materials presented are sound motion pictures, making the device particularly adaptable to the automated presentation of demonstrational sequences for both conceptual and procedural training. The device shown here differs from other sound motion pictures in the nature of its pre-programmable start-and-stop characteristics and also in the fact that, because its endless film loop is contained in a pre-threaded magazine, it can present any desired set of film sequences from a sizable library of magazines at virtually a moment's notice, with no need for threading or handling the film by the user. Thus, the mechanics of projection are removed from the attention of the instructor or teaching assistant, being turned over to the individual student and to the machine itself.

The sketch in Figure VI depicts the utilization of such a machine as an automated demonstrator for teaching procedural skills in the laboratory or in apprentice training of technicians on the job. The demonstrational segments to be performed by the student can be tailored in length to his ability and prior experience, so that he can perform each step and check his work as he proceeds against a picture that remains on the screen until he is ready to go on with the next step. Thus, the trainee can learn while he works and may perform useful work while he learns.

The general availability of such techniques of instruction in public schools must obviously await the development of an automatic projector, probably one with prismatic optics, that can be mass-produced at a very low unit cost. Such a projector has been for some years, and still is, one of the tools most needed for effective use of films in practical instruction.

Such step-by-step visual demonstration devices are not well suited, however, for teaching certain kinds of paced sequential performances such as keyboard skills. Some very interesting and ingenious teaching machines for teaching such skills (e.g., card punching) have been developed by Pask (1958a). These devices also utilize visual prompts, but their most interesting properties lie in the kind of electronically mediated feedback they provide to the student. This involves changes in speed requirements and variation of the practice materials presented in the program, depending on the student's current speed and the kinds of errors he makes. Pask (1960) also envisages the utilization of similar properties for other kinds of teaching and even of "adaptive" teaching machines that "learn" how best to adapt to the idiosyncrasies of an individual student.

Especially intriguing is the possibility that such devices may permit almost infinite variation in programs, which the machines

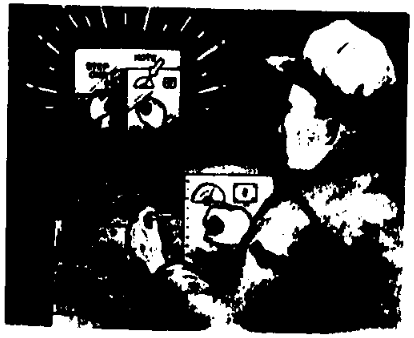


FIG. VI. Use of special step-by-step film projection as a procedural training aid. A technician watching a step-by-step procedure presented in short sequences is able to copy lengthy and exacting procedures accurately with minimum training.

themselves may, in effect, construct, to adapt to the behavior of each student. The psychological advantages and economic utility of these more sophisticated electronic machines as compared with simpler devices are interesting questions for speculation.

One of the general problems in the design of all forms of teaching machines is the question of program variation. In the Skinner and Briggs machines the program is fixed, except for drop out of mastered items, for some 30 frames or items, at the end of which the student may, depending on his performance and needs, go on to the next disk or go back to an earlier set of material to remedy some deficiencies in his prerequisite knowledge, or he may skip ahead to more advanced material.

With the magazine projector shown in Figure VI, the material may similarly be varied in units of five minutes or less, simply by selecting and inserting a new magazine. In some instances, even greater flexibility may be desired—so that the next problem or question presented to the student is governed by his response to each preceding question. This is particularly likely to be desirable in teaching problem-solving or diagnostic skills and may be useful in other teaching as well. At least three methods have been used thus far to achieve this kind of flexibility. One is the use of a digital computer (see Rath, Anderson, and Brainerd, 1959) which automatically adjusts the problem difficulty to each response that the student makes. This may seem like a fantastic degree of instrumentation, but it actually has practical possibilities for future development.

A second method uses microfilm projection equipment, such as is shown in Figure VII.

This device (see Crowder, 1959, 1960b) was developed for teaching electronic troubleshooting to Air Force technicians. It provides for a variable program in which not only the pacing but also the sequencing of instructional items can be varied, almost without limit, to adapt to qualitative differences in the requirements of individual students as indicated by their responses to previous items. This machine, and some similar ones developed experimentally in the former Air Force laboratory in Denver under the direction of R. M. Gagné, can select any sequence of questions or problems from several hundred items stored on microfilm, projecting each item in an order that is determined by the response of the student to the preceding item or items. This particular device was designed as a tool for research in methods of teaching concepts and discriminations which Air Force technicians need to master for troubleshooting complex electronic equipment systems. However, its basic principles are adaptable to the teaching of a wide variety of subjects, particularly those that require the acquisition of diagnostic or problem-solving skills.

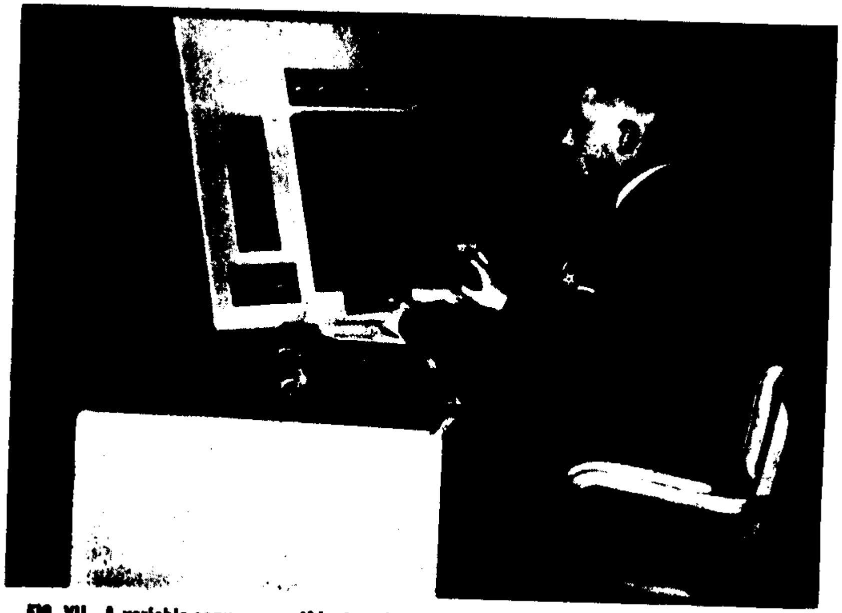


FIG. VII. A variable-sequence self-instructional device. The program of materials presented on the screen can be arranged in such a way that the presentation of new items of information depends upon the student's own performance.

A highly simplified, though less flexible and less automatic, variant of this device has been developed by Crowder (1959) in the form of a so-called "scrambled book," in which the pages are in random order. With this "paper teaching machine," the page that the student is told to turn to next is contingent on the answer he gives to the problem or question presented on the current page.

The field of teaching machines and related devices has recently been expanding so rapidly that published information has lagged far behind the work accomplished and in progress. Major products of the extensive work of Skinner and his co-workers up to last year, of the antecedent work of Pressey and his students, and a growing number of other investigators have, however, been described in a number of unpublished papers as well as in published articles. [Many of these are printed, some of them for the first time, in the later sections of this book.]

Christmas Past, Present, and Future: A Review and Preview

ROBERT GLASER

IN DISCUSSING the past, present, and future of teaching machines and programmed learning, the writer would like to remind educational psychologists of Ebenezer Scrooge in Charles Dickens' *A Christmas Carol.* Scrooge, it will be recalled, was a rigid, miserly man who did not appreciate the potential of liberal reinforcement. It was only when his history passed before his eyes in the form of three ghosts that he became aware of his past mistakes, present circumstances, and future prospects. It will further be remembered that Scrooge emerged a changed man. With this parable in mind, let us consider the articles by Pressey (whose forward-looking early articles might remind us of the ghost of Christmas Past); by Skinner (whose recent provocative article may take the role of Christmas Present); and by Ramo (whose speculations might be thought descriptive of Christmas To Come).

These articles represent key publications in the present flurry of excitement among psychologists and educators concerning the "teaching-machine" movement. The growing enthusiasm of many psychologists might well have been anticipated: The actuality of being able to arrange conditions so that human behavior can be brought from one level of performance to another level under the guidance of specifiable techniques is a very reinforcing event.

Psychology's pioneer in automated teaching is S. L. Pressey. Dr. Pressey's first papers present a device, an early version of which he exhibited and discussed at the meetings of the American Psychological Association in 1924 and 1925. In two initial papers reviewed here, Pressey discusses a "simple apparatus which gives and scores tests—and teaches." He points out that

---

This article appeared in the January 1960 issue of *Contemporary Psychology.* It reviewed certain key publications on teaching machines that appear later in this book: the articles by S. L. Pressey (1926, 1927), the article by B. F. Skinner (1958c), and the article by Simon Ramo (1957).

laborsaving devices are quite feasible in education and sees no reason why education should not be run as efficiently as any large-scale undertaking in this country. Rather than stultifying education, such mechanical aids should free the teacher from unnecessary burdens and leave her free "for those inspirational and thought-stimulating activities which are, presumably, the real function of the teacher."

The device Pressey presents contains a window through which a multiple-choice item appears. The subject has before him four keys and presses the key corresponding to his choice of an answer. Only when the correct key is pressed does a new question appear in the window. The apparatus can be adjusted so that, after a criterion of successive correct responses to an item, the machine does not present that item again to the student. Finally, after every item has been answered correctly to the specified criterion, the apparatus releases a small coupon or piece of candy indicative of the fact that the exercise has been mastered.

Pressey built his device to function in accordance with existing knowledge about the learning process. He wrote:

> . . . the "law of recency" operates to establish the correct answer in the mind of the learner, since always the last answer chosen is the right answer. The correct response must almost inevitably be the most frequent, since the correct response is the only response by which the learner can go on to the next question, and since whenever a wrong response is made, it must be compensated for by a further correct reaction. The "law of exercise" is thus automatically made to function to establish the right response. Since the learner can progress only by making the right reaction, is penalized every time he makes a wrong answer by being required to answer the question one more time, and is rewarded for two consecutive right responses by the elimination of that question, the "law of effect" is constantly operating, to further the learning. Finally, certain fundamental requirements of efficiency in learning are met. The learner is instantly informed as to the correctness of each response he makes (does not have to wait until his paper is corrected by the teacher). His progress is made evident to him by the progressive elimination of items. And—most important of all—there is that individual and exact adjustment to difficulty mentioned at the beginning of the paper, by which wasteful overlearning is avoided and each item returned to until the learner has mastered it. (Pressey, 1927).

Comment on Pressey's work is contained in Skinner's article (1958c), written a generation later. In this paper Skinner noted that the "industrial revolution in education" which Pressey envisioned did not come about. Skinner attributes this to cultural inertia. The world of education was not ready. The technological attitude encouraged by varieties of audio-visual aids, film projectors, television sets, phonographs, and tape recorders had not blossomed. Furthermore, the science of learning had not brought us to an understanding of the learning process such as we have today. Much of our knowledge, Skinner points out, has come from what has been learned by studying lower organisms. A technology of education based on what we now know is directly applicable to human learning, and the task ahead for the development of this technology calls for ingenious application of much of what we already know.

In order to arrange the conditions which change the behavior of lower organisms, we have learned that we often need to build elaborate apparatus that permits the control of variables of which learning is a function. It is reasonable to assume, then, that even more subtle manipulation of conditions is required to change the behavior of the human organism, much more subtle than can be handled by a teacher in a classroom. Appropriate instrumentation, that can be called a teaching machine, might be employed to present and schedule the complex contingencies required.

In contrast to Pressey's, Skinner's machines require that the student compose his response rather than select it from a set of alternatives. Further, in acquiring the behavior being learned, the student must pass through a carefully designed sequence of steps. Each step must be so small that it can always be taken; and as it is taken, it moves the student closer to the end-product behavior. These requirements are generated by what has been discovered about the shaping of operant behavior. A feeling for the operation of machines of this kind is conveyed by the following paragraph in Skinner's article:

Exit Behavior

> The machine itself, of course, does not teach. It simply brings the student into contact with the person who composed the material it presents. It is a laborsaving device because it can bring one programmer into contact with an indefinite number of students. This may suggest mass production, but the effect upon each student is surprisingly like that of a private tutor. The comparison holds in several respects. (a) There is

> a constant interchange between program and student. Unlike lectures, textbooks, and the usual audio-visual aids, the machine induces sustained activity. The student is always alert and busy. (b) Like a good tutor the machine insists that a given point be thoroughly understood, either frame by frame or set by set, before the student moves on. Lectures, textbooks, and their mechanized equivalents, on the other hand, proceed without making sure that the student understands and easily leave him behind. (c) Like a good tutor the machine presents just that material for which the student is ready. It asks him to take only that step which he is at the moment best equipped and most likely to take. (d) Like a skillful tutor the machine helps the student to come up with the right answer. It does this in part through the orderly construction of the program and in part with techniques of hinting, prompting, suggesting, and so on. . . . (e) Lastly, of course, the machine, like the private tutor, reinforces the student for every correct response, using this immediate feedback not only to shape his behavior most efficiently but to maintain it in strength in a manner which the layman would describe as "holding the student's interest." (Skinner, 1958c).

Pointing out that the success of a teaching machine depends upon the material used in it, Skinner devotes a large portion of his article to a discussion of the task of "programming" a given subject. He presents illustrative materials in spelling and high-school physics, describes their characteristics, and presents some rudimentary principles for programming techniques. Skinner discusses the question of maximizing success and minimizing failure and errors in the course of programmed learning, and also the question of how machine instruction will influence the role of the teacher. (The reviewer wishes to comment later on some of these matters.)

Skinner's article is provocative and challenging not only for education, but for the psychologist in the laboratory. The experimental analysis of the properties of verbal learning sequences is an old problem to which the programming requirements of such sequences might bring some new insights. Stimulated by the recent work of Skinner and his associates, a number of research and development efforts are in progress. Some individuals are concerned with building learning programs for particular subject matters and discovering and controlling the effective variables in a learning sequence. Other psychologists have been concerned with instrumentation and the characteristics of teaching machines as they influence stimulus presentation and the form of the learner's response. Others are concerned with experiments to study the effect of standard learning variables, like massed versus spaced practice and so forth, in the teaching machine context.

In a 1957 article Simon Ramo, a distinguished electronics engineer and designer of missiles, speculates about the schools of the future. His concern is that, given the rate of present technological and scientific advances, education will not be able to keep up with the pace of the new technical society; and he feels that it is the obligation of individuals who are engaged in engineering to apply themselves to help the process of education. He pictures a high school of the future in which the student, upon completion of registration, receives a specially stamped small plate (like a charga-plate) that identifies him and his course of study. Introduction of this plate into a machine makes available the entire record and progress of the student; it also presents to the student his program for the coming term.

A typical school day consists of a number of sessions, some of which are spent in rooms with other students and a teacher and some of which are spent with a machine. Sometimes a human operator is present with the machine and sometimes not. After a short period with a human teacher and after a motion-picture lecture, the student goes to the machine room. The student places his identification plate in a slot, his attendance is recorded, and the machine is connected with the master records machine. The machine then proceeds to tutor the student. Some machines present questions of the multiple-choice type. The progress of the student is then recorded by a master scheduling device. This sets the student up for other machines which are adjusted to his special needs and teaches him in Socratic-like question-and-answer fashion. Information from weeks of machine operation then form a record which the teacher uses to determine his further course of study. "Ultimately," writes Ramo, "with the proper cooperation between experts in education, expert teachers, experts [in specific subject matters], and experts in engineering these automatic systems, we can evolve that high level of match between the human teacher and the machine that we seek in that improved high school."

X

Ramo briefly mentions the use of machines for teaching the laboratory aspects of science, English composition, and written and spoken foreign languages. This new high school would require administrators and clerks not necessarily trained in education. There would be highly skilled teachers, for conventional teaching would still be a substantial part of the operation. In addition, however, the teachers would work closely with subject-matter experts and education engineers (a new profession). The requirements of this new education would create a substantial new industry concerned with creating these educational machines. (The 1959 meetings of the American Psychological Association contained several exhibits by teaching-machine manufacturers.)

> The high school becomes partially transformed into a center run by administrators and clerks, with a minimum of the routine assigned to the teaching staff. The teaching staff is elevated to a role that uses the highest intelligence and skills. A smaller number of teachers make possible the education of a larger number of pupils. The creation of education material moves partially out into industry, which goes into the education business in partnership with the educators.

It is the opinion of the reviewer that when this close partnership comes about it would be desirable if the science of psychology could offer technical assistance. Indeed, selling instructional aids and materials to our school systems is already a major business enterprise. The industries concerned should seriously consider the increased sponsorship of research in learning and teaching.

In essence, these papers recommend the application of what we know about learning to educational psychology, an "applied psychology of learning." This is an interesting point, for it implies that, for the most part, educational psychology to date has not fruitfully interacted with its basic-science counterpart, the science of learning. This essentially has been the true state of affairs. Experimental psychologists and educational psychologists are trained in different academic worlds and in different universes of discourse. As a result, educational psychology, and the teaching practices it generates have not been closely nurtured by their mother science. Old fashioned practices, long de-emphasized, are still applied and the potential of new developments is not recognized.

However, the signs of the times indicate that "togetherness" is important and essential. In recent years more experimentally and learning-theory-oriented psychologists have been working on education and training problems. The number of psychologists employed by the government and by private organizations to work on education and training in the military field has increased many-fold. The approaches to training problems made by these individuals and their reported records of success and failure in the application of the methods and knowledge of their science have important implications for education in general.

In the universities, educators and experimental psychologists are talking to each other more than ever before and joint appointments are being encouraged. The Office of Education of the U. S. Department of Health, Education, and Welfare is reinforcing this togetherness with research support. Stimulated by the apparent potential for educational practice that "programmed learning" and "automated teaching" methods offer, a number of smaller colleges are in the midst of experimenting with radical curriculum changes—not so much as to subject-matter content, but with respect to the teaching of it. The articles by Pressey, Skinner, and Ramo encourage this *rapprochement* between the "science of learning and the art of teaching." (Skinner, 1954.)

The reviewer must now comment further on several points.

1. *Programming material.* Skinner wisely emphasizes that the success of a teaching machine depends on the material used in it. The essential task set for the programmer is to evoke specific forms of behavior and, through differential reinforcement, to bring them under the control of specific stimuli. As a student goes through a learning program, certain of his responses must be strengthened and stimulus control transferred from one situation to another. How does one go about doing this? Some techniques have been suggested by Skinner and others. However, the reviewer and his colleagues, in the course of constructing programmed learning sequences, have been impressed with the artistic nature of the enterprise. The fact that it is a difficult and aversive task to program material and a much easier task to build the accompanying hardware is indicated by the fact that at the present time, machines outnumber programs by a large factor. We are in the situation of having

shells without innards. Once programmed learning sequences are prepared in various subject matters that teach better than existing teaching methods, an educational revolution may not be far behind. Once the airplane flies, the air age cannot be ignored; but it cannot fly without the inside workings.

Learning programs based upon even our present knowledge can scarcely fail to be an improvement over anachronistic methods of teaching certain subjects by lecturing to large classes with little close control of student behavior in acquiring the subject matter. The task indicated is a de-emphasis on the hardware and a concerted attack on the programming of materials and the development of specific principles of programming learning sequences. It must be said, however, that the hardware has its place. It provides a means of manipulating the display characteristics of the reinforcement contingencies employed by the device. Finally it offers "tangibility" and the advantages of automation to the practicing educator. Methods implemented by machine are easier to sell than methods alone.

2. *Defining the field.* Skinner also states that in programming materials the first step is to define the field. This is no trivial concern; the problem is the specification of the behavioral end-products of the learning process. With the control offered by programmed learning sequences, fairly precise statements must be made of the form of the knowledge skills that are to be learned. In teaching mathematics, for example, do we want the student to solve problems, apply his knowledge to new situations, prove theorems, or all three of these? The construction of the form of the learning programs can differ radically as a function of the criterion behavior chosen. Heady with the possibility of close control of the learning of students in our school systems, the reviewer is tempted to make a tempered Watson-like statement: "Specify the behavior you want acquired, and I can get the student to learn it more effectively than ever before." What, however, is the best way to specify this knowledge in order to facilitate the development of effectively controlling learning sequences? A concerted attack on the behavioral specification of educational objectives is required. It is interesting to speculate about how this might influence the structure and organization of our various fields of knowledge.

X

3. *Feedback on effectiveness.* A programmed learning sequence provides constant feedback about its effectiveness. Does it teach the criterion behavior? Does it lead the student in the direction required to master the subject matter under consideration? It seems apparent to the reviewer that programmed material can teach many things more effectively than they are taught at the present time. In addition, these new techniques can be evaluated more severely than ever before. When the question is asked, "Does this textbook or lecture teach effectively?" the question is answered by the authority of the textbook writer, editor, or lecturer. However, when this question is asked of programmed material, it is answered by detailed analysis of the behavior of the student (similar to the detailed analysis of test items). This behavioral orientation of the learning program permits great improvement in the course of the development of programmed materials

Finally, if the reviewer's (and other psychologists') enthusiasm for the potentialities of an applied psychology of learning has sounded overzealous, it is because now seems to be the time for such zealotry, or else another 35 Christmas seasons—like the years since Pressey's APA exhibit—will pass and no educational revolution will have occurred. Perhaps schools of education are too exclusively concerned with turning out teachers with a well-rounded view of the general problems of education and are not sufficiently concerned with practice in the application of knowledge from the modern science of learning. On the other hand, perhaps psychologists are too exclusively concerned with turning students of experimental psychology into future Einsteins and into ultra-careful experimentalists, and need to temper this with the production of some Edisons and inventive applied scientists. At least let us try to apply effectively as much as we know; it might be enough to make a difference. The stimulus provided by the recent concerns with "teaching machines" and "programmed learning" makes this result seem likely.

PART II

Pressey's Self-Instructional Test-Scoring Devices

IN THE early 1920's the modern objective test suggested to Sidney L. Pressey of The Ohio State University the possibility of using a mechanical device for administering and scoring tests. He further saw the potential of such a device for performing certain routine functions of a teacher in drill and recitation if the device was arranged to supply the student with immediate information concerning the correctness of each response. It seemed evident to Pressey that this immediate knowledge of results could afford a more effective means of teaching than commonly used classroom procedures. He also pointed out that learning with a mechanical device could implement several principles, or "laws," of effective learning.

Pressey's implementation of these basic notions is described in the first two of the papers which follow, published in 1926 and 1927. In these papers he describes mechanical devices for presenting multiple-choice questions and for giving the student immediate feedback concerning the correctness or incorrectness of his answers. He points out that not only are these devices laborsaving for the teacher but also that they make it possible for the learning of the individual student to be regulated in accordance with principles of learning derived from detailed study of the learning process. Although various "educational devices" had been previously built, Pressey seems to have been the first person to make a systematic attempt to embody into a device the laws of learning, as formulated at that time. His effort to do this is evident in both papers. These two basic papers are reprinted here in full, with the exception of some mechanical details which Dr. Pressey feels are no longer of current interest. Dr. Pressey has also provided photographs of the two devices described in these early articles, to replace the line drawings originally used.

Following the presentation of his basic notions in the 1926 and 1927 papers, Pressey and his students at The Ohio State University continued to develop additional devices and to try these out experimentally with a variety of course materials. Six further papers stemming from this work are included here, and several additional papers are abstracted in Appendix I. The selection of papers has followed the recommendations of Dr. Pressey, who has also kindly undertaken to edit these later papers for this volume. In doing this, he has deleted or abridged various portions which he feels are not of current interest but has retained the key material that exemplifies the major phases of the work.

Pressey's early hopes for the future of education based on a technological revolution in educational procedures and instrumentation are further developed in his third article (1932). He points out: "Education is the one major activity in this country which is still in the crude handicraft stage." A significant implication of this article, in which two further innovations are also described, is the importance of new forms of instrumentation as tools through which the science of learning can be advanced.

An early, simplified variant of Pressey's multiple-choice response apparatus was the self-scoring "chemo card" devised by Pressey's former student H. J. Peterson and his brother, J. C. Peterson, and described by the latter in 1930 (see Appendix I for an abstract of this paper). This was used in an early experimental comparison to assess the value of the immediate knowledge of test results which it provided. The fourth of the papers reprinted here is a 1931 report on this experiment by J. C. Peterson. This is the first of a series of experimental studies suggesting the potential effectiveness of devices based on Pressey's concepts.

A fifth paper in this series represents a second early investigation reported by J. K. Little in 1934. This study, conducted as a doctoral dissertation under Pressey's direction, employed two devices for providing prompt knowledge of results to students—a mechanical test-scoring device and one of the Pressey multiple-choice drill machines. The results again suggest the potential benefits to be gained by the use of such instructional devices.

After these early publications there is a gap of about 15 years before the next series of papers growing out of Pressey's work appeared in the late forties and early fifties. The first of these papers, published by G. W. Angell and M. E. Troyer in 1948, mentions the use of devices by the Armed Forces during World War II for giving immediate knowledge of results, and describes a simple punchboard apparatus, variants of which were subseqently used in several studies at The Ohio State University. These include studies by Briggs (1949), Jensen (1949), Jones and Sawyer (1949), Jones (1954), and Severin (1955), abstracts of which are included in Appendix I. Several of these abstracts have been provided by Dr. Pressey,[1] under whose direction the studies were performed.

Pressey's 1950 paper provides a summary of the series of studies done under his direction during the late forties, using the punchboard as a self-instructional device. An abridged version of this paper, prepared by Dr. Pressey for the present volume, discusses the use of the punchboard in a variety of educational applications. These include the use of the device as a primary form of instruction as well as an adjunct to other teaching and its use with selected groups of superior students.

The final paper in Part II reports experiments by A. L. Stephens, another of Pressey's students. An interesting aspect of this study was the use of a rather wide range of materials, including nonsense syllables, Russian and English vocabulary, and college course material in psychology. For these materials, Stephens compared test procedures that provided varying forms of knowledge of results.

Collectively, these studies represent a concerted effort both in device development and in attempts to try out and to assess the contribution of using the devices in various ways. Although the methodology and experimental controls employed do not in all cases provide a basis for definitive conclusions, these papers represent an important and persistent development effort in the use of self-instructional devices. Basically, the objective was to provide more effective learning through instrumentation which maximizes active student response and immediate knowledge of results.

---

[1] A recent reaction by Pressey to current developments in teaching machine and programming concepts is presented in Part V, p. 497-505.

A Simple Apparatus Which Gives Tests and Scores – and Teaches

S. L. PRESSEY

For a number of years the writer has had it in mind that a simple machine for automatic testing of intelligence or information was entirely within the realm of possibility. The modern objective test, with its definite systemization of procedure and objectivity of scoring, naturally suggests such a development. Further, even with the modern objective test the burden of scoring (with the present very extensive use of such tests) is nevertheless great enough to make insistent the need for laborsaving devices in such work.

The writer has also felt that the procedures in mastery of drill and informational material were in many instances simple and definite enough to permit handling of much routine teaching by mechanical means. The average teacher is woefully burdened by such routine of drill and information-fixing. It would seem highly desirable to lift from her shoulders as much as possible of this burden and make her freer for those inspirational and thought-stimulating activities which are, presumably, the real function of the teacher.

The present brief paper deals with a simple apparatus which automatically gives *and* scores a test, and which will also, automatically, teach—and teach informational and drill material more efficiently, in certain respects, than the "human machine."

USING THE APPARATUS

The apparatus (Figure I) is about the size of an ordinary portable typewriter—though much simpler.[1] . . . The person

---

This article is reprinted from *School and Society*, Vol. 23, No. 586, March 20, 1926. Figure I in the original article was a line drawing which has been replaced here with a photograph furnished by Dr. Pressey.

[1] The device was exhibited, and described in a brief paper, at the meetings of the American Psychological Association in Washington in December 1924 and exhibited in improved form at the 1925 meetings of the same organization.

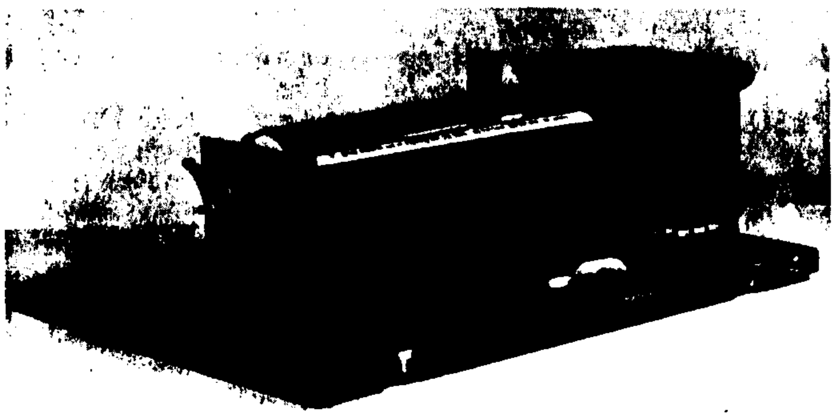


FIG. I. A self-scoring multiple-choice apparatus exhibited in 1924-25

who is using the machine finds presented to him in a little window a typewritten or mimeographed question of the ordinary selective-answer type—for instance:

> To help the poor debtors of England, James Oglethorpe founded the colony of (1) Connecticut, (2) Delaware, (3) Maryland, (4) Georgia.

To one side of the apparatus are four keys.[2] Suppose now that the person taking the test considers Answer 4 to be the correct answer. He then presses Key 4 and so indicates his reply to the question. The pressing of the key operates to turn up a new question, to which the subject responds in the same fashion. The apparatus counts the number of his correct responses on a little counter to the back of the machine . . . . All the person taking the test has to do, then, is to read each question as it appears and press a key to indicate his answer. And the labor of the person giving and scoring the test is confined simply to slipping the test sheet into the device at the beginning (this is done exactly as one slips a sheet of paper into a typewriter), and noting on the counter the total score, after the subject has finished.[3]

The above paragraph describes the operation of the apparatus if it is being used simply to test [Set 1]. If it is to be used also to teach [Set 2] then a little lever to the back is raised. This automatically shifts the mechanism so that a new question is not rolled up until the correct answer to the question to which the subject is responding is found. However, the counter counts all tries.

It should be emphasized that, for most purposes, this second set is by all odds the most valuable and interesting. With this second set the device is exceptionally valuable for testing, since it is possible for the subject to make more than one mistake on a question—a feature which is, so far as the writer knows, entirely unique and which appears decidedly to increase the significance of the score. However, in the way in which it functions at the same time as an "automatic teacher" the device is still more unusual. It tells the subject at once when he makes a mistake (there is no waiting several days, until a corrected paper is returned, before he knows where he is right and where wrong). It keeps each question on which he makes an error before him until he finds the right answer; he *must* get the correct answer to each question before he can go on to the next. When he does give the right answer, the apparatus informs him immediately to that effect. If he runs the material through the little machine again, it measures for him his progress in mastery of the topics dealt with. In short, the apparatus provides in very interesting ways for efficient learning.[4]

---

[2] By using only two keys the apparatus can, of course, be employed to handle the "true-false" type of test.

[3] In the demonstration apparatus at present being used, there is an additional attachment which drops into a little container before the subject, a small piece of candy, if he makes the number of correct responses for which the experimenter has set the "reward dial." With this attachment, the apparatus thus does one thing further; it automatically rewards the subject as soon as he reaches whatever goal the person giving the test may set for him.

[4] The somewhat astounding way in which the functioning of the apparatus seems to fit in with the so-called "laws of learning" deserves mention in this connection. The "law of recency" operates to establish the correct answer in the mind of the subject, since it is always the *last* answer which is the right one. The "law of frequency" also cooperates; by chance, the right response tends to be made most often, since it is the *only* response by which the subject can go on to the next question. Further, with the addition of a simple attachment, the apparatus will present the subject with a piece of candy or other reward upon his making any given score for which the experimenter may have set the device; that is, the "law of effect" also can be made, automatically, to aid in the establishing of the right answer.

THE MECHANISM

The general nature of the mechanism is shown in the schematic drawing below (Figure II), which should be studied in relation to Figure I. As shown in Figure I, the paper on which the test question appears is carried by a typewriter platen—and in exactly the same manner as paper is carried in a typewriter. The keys which the subject presses in taking the test appear in both figures. The "inner workings," which bring about the various operations mentioned in the previous section of the paper, are shown in Figure II.

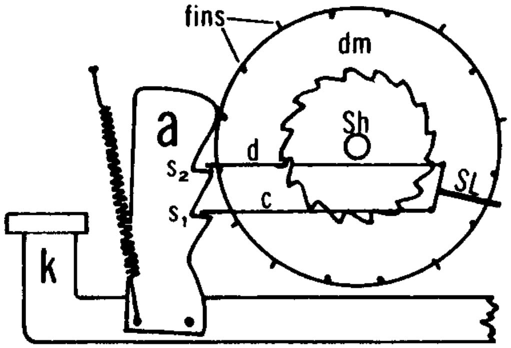


FIG. II. A schematic drawing of the internal mechanism of the device shown in Fig. I.

The fundamental principle is simple. The shaft, on which the platen is carried, carries also a drum (dm).[5] On this drum are rows of little projecting fins—one row for each question and four fins in each row (one for each possible answer). These fins can easily, with a fingernail, be turned in, so that they are almost flush with the drum or turned out so that they project slightly. What the experimenter does in setting the apparatus for a particular test sheet is simply to "gear" this drum to the test by turning in the fin corresponding to each right answer.

When now the key (k) is pressed (Figure II shows the key pressed down), the shoulders s1 and s2 on the activating upright (a) will, in the return movement, engage with the ends of the platen-drive arm (d) and counter-drive arm (c). That is, the platen will roll up a new question, and the counter (an ordinary Vidor counter, not shown in the figures) will count. . . . All this will happen if the "correct" key was pressed. Suppose the subject had pressed the wrong key. The activating upright of that key would have been held back by a fin—back far enough so that the shoulder s2 would have missed the end of the platen-drive arm (d). However, shoulder s1 would have still engaged the end of the counter-drive arm (c). Thus no new question would be turned up, but the counter would count a try (the description is for operation with the more important Set 2, for testing *and* teaching).

Suppose now the experimenter wishes to shift the apparatus to Set 1. He simply raises the shift lever (SL). This pushes the platen-drive arm (d) forward slightly and draws back the counter-drive arm (c). The result is that even when a fin holds back the activating upright (a), shoulder s1 will engage and the drum will be turned; but the counter-drive arm (a) will not engage. That is, the apparatus will count only rights but will always turn up a new question.

The exact details of the construction are of no great importance here; as a matter of fact, a number of alternative forms are possible.[6] The important point here is that a mechanism as comparatively simple[7] as that indicated by the illustrations (quite as simple as the ordinary alarm clock) will do the things mentioned under "Using the Apparatus."

---

[5] These operating parts are all within the square casing shown in Figure I at the right end of the apparatus from which the keys project.

[6] With electrical operation, there are obviously a great variety of possibilities (for description of a "continuous choice-reaction apparatus" embodying somewhat the same essential principle, developed by the writer some 10 years ago, see the *American Journal of Psychology*, Vol. 32, p. 347 ff.). The writer is now working upon a drill apparatus which *will omit a question from further presentation as soon as the subject has obtained the correct answer twice in succession;* there is thus no excessive overlearning, and there is a concentration of effort on each problem in proportion to its difficulty. As a matter of fact, no teacher can thus nicely adjust to the detail of each child's needs. The device should be uniquely valuable in this respect.

[7] The only essential feature not included in the illustrations and description is the very simple device for preventing the subject from pressing more than one key at once.

LABORSAVING DEVICES IN EDUCATION

Now, briefly, as to the larger significance of the device. It raises the fundamental question as to whether laborsaving mechanisms may not be used in education as in other work. The writer has seen somewhere the statement that education is at present the most inefficiently carried on of any large-scale undertaking in this country. Insofar as this is true, it is due in large measure, the writer believes, to lack of any consistent application of *ingenuity* to the teacher's problems. Education is, of course, almost always conservative, and, comparatively, a teacher's time is cheap. The result is that there are at present many things now done in our schools and colleges in very unnecessarily labored and enthusiasm-killing fashion. The writer is convinced that mechanical aids are possible which would do much to relieve the situation. These aids would probably do their particular work better (just as a calculating machine is more accurate than the old-time bank clerk). More important —they would leave the teacher more free for her most important work, for developing in her pupils fine enthusiasms, clear thinking, and high ideals.[8]

Now, an illustration to show, concretely, what might be done. The writer is in immediate charge of a first course in educational psychology which is handled by a number of instructors. In this course it has been the custom to give each week a short quiz—usually of the objective-answer type. These quizzes take about five hours per section to score. There is an average of five sections per quarter for the four quarters of the year. The total time cost of scoring these tests is thus for the year about 1000 hours, or 125 eight-hour days—about five months (21 weeks) of one person's full time. Thirty-five of the above-described simple machines would have saved practically all this labor. Further, mistakes would be eliminated. The students would profit much more by the quizzes, since (with the second set of the apparatus) they would have, always, to obtain the right answer to each question. As a result of this last fact, little if any time would need to be given in class to going over the questions after the quiz. And the instructors could do very much more than at present in the way of class discussion, special help on difficulties, and so on. That is, in actual practice such mechanical aids would seem likely to prove of real, and by no means negligible, value.

---

[8] The writer's suggestions may very likely provoke some sentimentalists to an outcry against "education by machine." What the writer is urging is a freeing of the teacher from the mechanical tasks of her profession—the burden of paper work and routine drill—so that she may be a real teacher, not largely a clerical worker.

SUMMARY

1. It is pointed out that objective tests naturally suggest the possibility of a simple mechanism for testing. There are also some reasons for supposing that some of the teaching of drill material might be done by machine.
2. An apparatus is described which gives and scores tests and informs the subject with regard to the right answers (an attachment will reward the subject after any given number of right answers has been made).
3. It is emphasized that teachers are now heavily burdened with routine and clerical tasks that might well be handled mechanically—thus freeing the teacher for much more real teaching, of the thought-stimulating and ideal-developing type, than is now possible.

A Machine for Automatic Teaching Of Drill Material

S. L. PRESSEY

IN A previous number of this journal (Pressey, 1926) the writer described a "simple apparatus which gives and scores tests—and teaches." In this article it was emphasized that labor-saving devices in education should be entirely possible and should, instead of mechanizing education, leave the teacher free of much burdensome routine so that she could do more real teaching. In accordance with this general concept the writer has for some time been working (as mentioned in the paper already referred to) upon "a drill apparatus which will omit a question from further presentation as soon as the subject has obtained the correct answer twice in succession." As was indicated in this previous paper, such a device should prevent excessive overlearning and bring about a concentration of effort on each problem in proportion to its difficulty; and it was pointed out that such a device should adjust in these matters with a nicety which no teacher could well equal. The present paper describes this drill machine.

HOW THE APPARATUS APPEARS TO THE LEARNER

Figure I shows the way the apparatus appears to the subject.

Through the window in the casing there appears a question of the selective [multiple-choice] answer type . . . . This question is typed on a typewriter-size sheet of paper, which is clipped onto a drum within the casing. To the right of the device are four keys. If the subject considers Answer A right, he presses the first key; if he thinks Answer B is right, he presses the second key—and so on. When he presses the right key, the drum revolves and turns up a new question; the question is kept before the subject, however, (the drum does not move), until the right answer *is* found. The subject thus continues reacting, the apparatus presenting the questions in order (the drum continuing to revolve) and goes through the series the second, third, or further number of times. However, after the series has been gone through twice, the apparatus revolves past those questions which on both first and second trials were answered at once—without the pressing of a wrong key. And as each item is learned to the point where two successive right answers are made, it is thus thrown out. Finally, after every item has been mastered to this point, the apparatus automatically stops and releases a small coupon, indicative of the fact that that exercise has been mastered.[1]

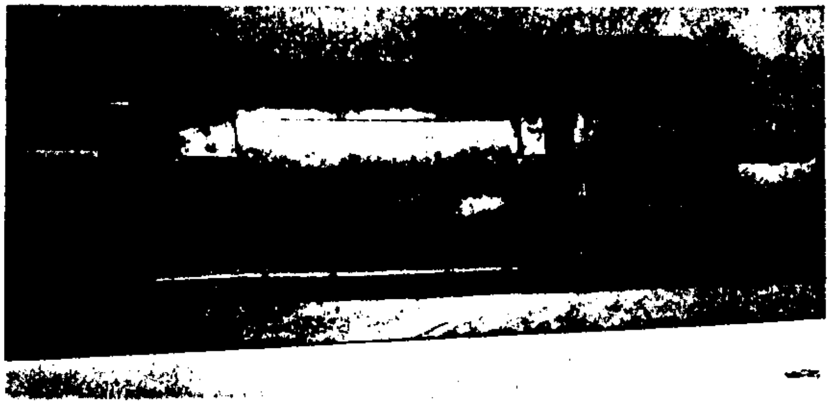


FIG. I. A multiple-choice device which omits items from further presentation once the student can consistently answer them correctly.

The apparatus thus keeps the subject at each question until he masters it and, when the mastery is achieved, takes the subject's time with that item no more. And as the learning progresses, the subject's attention is concentrated more and more on those items which are most difficult for him. There is just

---

This article is reprinted from *School and Society*, Vol. 25, No. 645, May 7, 1927. At the suggestion of the author, specific mechanical details of the apparatus described in the original article have been omitted here. Dr. Pressey has also furnished a photograph to replace the line drawing which appeared as Figure I in the original article.

[1] As a matter of fact, the apparatus is so adjusted that the user can set as a standard either two, three, or four successive right answers. The description is in terms of two successive right answers to make the statement more simple and because this standard seems to be the most satisfactory.

such repetition and centering of attention as most efficient teaching calls for. Then, once the learning is completed, the subject is automatically notified of this delightful fact—and the lesson is automatically ended.

HOW IT WORKS

Now, as to the "inside workings" of the device.

. . . .

When the correct key is pressed, the drum is released so that it can revolve. The revolution of the drum proceeds until a catch engages with a cog moving a bar inward, on a succeeding bar. When, however, a bar is thrown clean to the left as far as it will go, the cogs on that bar are thrown to a point where the catches will no longer operate on this bar. The result is that the drum will revolve past this question. And as the subject progresses in mastery of the material, he thus throws more and more bars to the left—out of the way of the catch mechanism—until finally, with elimination of the last and most difficult question, the drum is "cleared" and its rotation finally stopped.[2]

THE SIGNIFICANCE OF THE DEVICE

Now, briefly, as to the larger significance of the device. It should be noticed, first of all, that the apparatus is so built as to function, even more than the device described in the previous paper, in accordance with what is now known concerning the learning process. As with the first device, the "law of recency" operates to establish the correct answer in the mind of the learner, since always the last answer chosen [before the drum revolves to the next question] is the right answer. The correct response must almost inevitably be the most frequent, since the correct response is the only response by which the learner can go on to the next question; and since whenever a wrong response is made, it must be compensated for by a further correct reaction. The "law of exercise" is thus automatically made to function to establish the right response. Since the learner can progress only by making the right reaction, he is penalized every time he makes a wrong answer by being required to answer the question one more time, and is rewarded for two consecutive right responses by the elimination of that question, the "law of effect" is constantly operating to further the learning. Finally, certain fundamental requirements of efficiency in learning are met. The learner is instantly informed as to the correctness of each response he makes (does not have to wait until his paper is corrected by the teacher). His progress is made evident to him by the progressive elimination of items. And—most important of all—there is that individual and exact adjustment to difficulty mentioned at the beginning of the paper, by which wasteful overlearning is avoided and each item returned to until the learner has mastered it.

The apparatus, as described above, has certain questionable features, which are now being investigated.[3] But the merits or faults of this particular device are to the writer a somewhat secondary consideration. The important feature of the work here reported is the exemplification of the fact that machines can be built which meet, automatically, certain very important requirements of efficient teaching. This apparatus is a further evidence that laborsaving devices in education should be a possibility of the near future. Various devices should be possible.[4] The major purpose of this paper is to stimulate further work along these lines. As was emphasized in the previous paper, such devices should not operate to mechanize education in any unfortunate way (though educational sentimentalists may bring this charge). Rather they should free the teacher from much of the present-day drudgery of paper-grading, drill, and information-fixing[5]—should free her for real teaching of the inspirational.

---

[2] The locking mechanism consists essentially of a catch which is thrown up against the drum by the withdrawal of the last bar—and which at the same time kicks out a coupon or ticket as evidence that the task is done. This mechanism is supplementary to the major features of the device and is not shown in the figures.

[3] The first question is as to whether the selective-answer type of material, involving repeated presentation of wrong as well as right answers, is satisfactory material for instructional purposes. Theoretically it appears unfortunately possible that wrong association may be developed as well as the correct "bond"—though it might be argued that learning might involve not only the establishment of the right response but the avoidance of common wrong responses. The needful thing here is experimentally to compare learning "by machine" with learning by more usual methods; a graduate student is now making this comparison. If presentation of wrong answers is found unfortunate, then, in arithmetic at least, a different "hook-up" is possible—as will be mentioned shortly. The second question is whether reactions of the key-pressing type are so different from the reactions to be made in use of the materials being learned as to present a serious problem of transfer. This question also is being investigated. Briefly, it may be said that in much school learning there is some analogous problem of transfer; the objection is not believed serious.

SUMMARY

1. The paper reports an effort to develop an apparatus for teaching drill material which (a) should keep each question or problem before the learner until he finds the correct answer, (b) should inform him at once regarding the correctness of each response he makes, (c) should continue to put the subject through the series of questions until the entire lesson has been learned, but (d) should eliminate each question from consideration as the correct answer, for it has been mastered.
2. Such an apparatus is described (a) as it appears to the learner and (b) as to "inner workings."
3. It is reiterated that laborsaving devices should be possible in education. Such devices might well handle certain types of routine work even better than the teacher could. They should save the teacher's time and energy from such routine, so that she may do more real teaching of the ideal-developing and thought-stimulating type.

---

[4] Thus the writer has in mind first rough plans for a machine for drill in arithmetic, having the usual adding machine keyboard and having the problems either on a roll of paper perforated somewhat after the fashion of a player-piano roll or on perforated cards of the tabulating-machine type.

[5] And may very likely do this work (as has been suggested in the description of the apparatus with which this paper deals) in many respects better than the teacher could do it.

A Third and Fourth Contribution Toward the Coming "Industrial Revolution" in Education

S. L. PRESSEY

In two previous articles in *School and Society* the writer pointed out that education was the one major activity in this country which had thus far not systematically applied *ingenuity* to the solution of its problems; and he attempted to illustrate, by two sample devices, how simple machines might both score tests and—with certain types of material—actually teach. In this third and last article of the series, two further attempts at laborsaving devices in education will be briefly described and an attempt made to evaluate the present situation and future possibilities of work along these lines.[1]

THE GENERALIZED ANSWER UNIT AS THE NEXT IMPORTANT STEP IN "TEST FORM"

It is to be hoped that some day there will be written a history of the development of test form. . . . That history would doubtless begin with the introduction of checking and underlining as means of indicating the correct answer to a question. . . . To facilitate scoring, the scoring stencil and transparency were early introduced. . . . However, the average test has four serious faults as regards form: (a) It seems most unfortunate in these times of economic depression that elaborate test blanks costing usually $5 or more per hundred can be used only once—because the pupils mark upon them. (b) Since the pupil's answers are scattered through the booklet, there must be a separate scoring stencil or strip for each page. . . . (c) The

---

The original version of this article was published in *School and Society*, Vol. 36, No. 934, November 19, 1932. Dr. Pressey has edited the paper for inclusion in the present volume.

[1] Certain passages in the original article not now of interest and drawings giving details of apparatus have been here deleted by the author in re-editing for the present volume.

keeping of a detailed record of a test is almost impossible without keeping the bulky test booklets. (d) That natural extension of the test movement—the locally made objective examination—is greatly hampered. . . .

The solution of these difficulties which the writer wishes to present is very simple. . . . Others have experimented with the idea. However, these various trials of separate answer units seem to have attracted relatively little attention. And the possibilities of the separate answer unit in facilitating scoring seem to have been almost totally neglected. The problem was to produce the most simple and convenient possible answer and record unit and scoring scheme. The answer unit here suggested consists simply of an ordinary 3″ x 5″ card with 100 numbered answer boxes, 45 on the front and 55 on the back. . . . If the first question is a true-false question and the pupil considers the statement false, he marks "F" in answer box number one, and so on. . . . In scoring, the pupil's marked card is slipped into a double stencil, or "scoring corner," and the pupil's mark in each window is compared with the correct answer immediately to the left of the window. . . . The form of materials here reported is not suitable for all purposes. . . . But the writer ventures the prediction that some form of separate answer unit will in a few years be the usual procedure in testing.

A MACHINE FOR SCORING TESTS AND TABULATING BY ITEM

The second part of the paper reports a mechanical device for very rapidly tabulating by item as well as scoring and totaling tests. . . . The answer unit on which the student indicates his answers to questions consists of a strip of thin cardboard 1¾ inches wide by 11 inches long. Below a space at the top where the student can write his name are 30 rows, each of five 3/16 inch circles. The student indicates his answer to a question by punching with a hand-punch through the appropriate circle.

The scoring and tabulating machine is essentially simple in principle. The platen is a cast aluminum block about one inch thick, and in length and breadth the dimensions of the answer unit. In it are 150 holes corresponding to the circles on the answer unit. And in setting up the device for scoring a particular test, the operator inserts, in those holes which correspond with the holes in the answer unit which should be punched by the student, pins which project up from the platen by spring pressure—the springs being light enough that, if a right answer has not been punched, the cardboard will suffice to hold down that pin. In scoring a given answer strip, the operator first simply puts it on the platen; the side rails (between which the platen is fastened, these in turn being fastened to a base) and front end project above the platen sufficiently that positioning is easy.[2]

A little carriage is now run down the rails and back again. It carries a 10-gear pinion in mesh with a 30-unit counter wheel. Wherever a pin projects (that is, wherever a student has punched a right answer), the pinion is caught by this pin and rotated one cog, or one unit on the counter. On the down trip of the carriage the total number of right answers is thus counted and stamped at the bottom of the answer strip, on which the questions answered correctly are also marked by a dauber attached to a ratchet which keeps the pinion from turning more than one division for each pin. On the return trip the counter is cleared back to zero.

The method of tabulating by item is also simple. For each one of the 30 rows of holes on the platen and answer strip—that is, for each question—there is an item-counter unit consisting of a pawl and a 100-cog wheel fastened to a disk carrying on its edge numbers from 1 to 100 (read through a slit in a cover plate). Whenever the carriage runs over a right answer (that is, wherever a projecting pin trips the pinion) the carriage ratchet is jogged up, pushing the corresponding item-counter pawl into engagement with the next cog on the counter wheel. At the end of the return movement of the carriage, the yoke carrying all the item-counter pawls is shoved over just enough that each item counter, the pawl of which has been engaged by the stroke of the carriage ratchet, is turned one unit. And after the answer strips for a given class have been run through the machine, each item counter shows the number of students who answered that question correctly.

As will be seen, the essential construction of the device is simple. The action is very rapid (a page being scored and tabulated by item in about one second); the very great saving over hand-scoring and tabulation is very obvious. Various improvements are possible. In fact, it may very well be that some very different type of registration (perhaps a photoelectric cell) may be most satisfactory. But the device here reported at least demonstrates that a relatively simple mechanism permitting very rapid scoring of tests and tabulating by item is entirely feasible.

---

[2] The original paper included detailed drawings, reproduction of which here seems unnecessary for explication of the general nature of the device.

THE POSSIBILITY OF AN "INDUSTRIAL REVOLUTION" IN EDUCATION AND OF A DISTINCTIVE INSTRUMENTATION IN THE SOCIAL SCIENCES

In this series of three articles the writer has tried to exemplify the following points: (a) Great savings in material and labor, in the handling of the usual tests, can be effected by the use of separate answer units with appropriate special scoring devices; (b) tests can be automatically scored and results tabulated by item in one very rapid operation by machine; (c) it is entirely possible, by machine, not only to test a student but also to correct his mistakes, to teach him the right answers, and to automatically reward him for good scores (as illustrated in the first article);[3] (d) drill or teaching devices go still further, as by adjusting the amount of practice on each item to the extent to which the student has mastered that item (as described in the second article), are entirely feasible and suggest the possibility of relatively close control, by machine, of important features of the learning process.[4] The demonstration of these four propositions in terms of concrete working devices has constituted the writer's major interest during the past eight years, to which he has devoted himself all too completely. Any one of these undertakings, considered by itself, might seem only a bit of freak research. But it is hoped that all four taken together may be sensed as of some general significance. Two predictions are ventured:

1. Education is the one major activity in this country which is still in a crude handicraft stage. But the economic depression may here work beneficially, in that it may force the consideration of efficiency and the need for laborsaving devices in education. Education is a large-scale industry; it should use quantity production methods. This does not mean, in any unfortunate sense, the mechanization of education. It does mean freeing the teacher from the drudgeries of her work so that she may do more real teaching, giving to the pupil more adequate guidance in his learning. There may well be an "industrial revolution" in education. The ultimate results should be highly beneficial. Perhaps only by such means can universal education be made effective.
2. The advance of a science is closely dependent upon the development of instruments in that science. There has so far been relatively little development of instruments specifically for the very extensive yet analytical research typical both of modern educational investigation and also more generally of the social sciences. New instruments and materials, greatly facilitating research, may soon appear. There may then be sweeping research advances in these fields.[5]

---

[3] It may be mentioned here that mimeographed material for experiments in learning and memory, keyed to this apparatus, have been developed.

[4] As another historical aside, it may be remarked that the "educational toy" has a long history of potential significance in connection with the development of mechanical aids to teaching; the great trouble seems to have been that nobody took these "toys" seriously! But the Russians are now apparently using such devices. The writer is now making a collection of such toys.

[5] The writer regards the entire field touched upon in these studies as still almost wholly unexplored. Even in pencil-and-paper techniques there remain great possibilities. Thus one of the writer's former students, H. J. Peterson, is working on an invisible chemical, whereby when a pupil marks the right answer on an objective test, his mark will at once turn red. The pupil "workbooks" now so widely used in arithmetic, reading, and science may well bring about new paper-and-pencil techniques. And why not machines for drilling a pupil in arithmetic, giving and correcting his spelling lessons? Why not voting machines to gather data for the sociologist, psychologist? The problems of invention are relatively simple; with a little money and engineering resource, a great deal could easily be done. The writer is regretfully dropping further work on these problems [1932]. But he hopes that enough has been done to stimulate other workers.

The Value of Guidance in Reading for Information

J. C. PETERSON

THIS STUDY is an experimental investigation of the value of guidance in learning by means of a new automatic self-checking device, the Self-Instructor and Tester.[1]

In its simplest form this device consists of one or more sheets of multiple-choice questions whose correct and incorrect answers are treated with two matched moisture-sensitive inks which turn to contrasting colors when moistened. In using this device, the learner reads each question, chooses from its alternative answers the one he deems correct, and touches it with a strip of moistened felt. If the chosen answer is correct, it changes immediately to a predetermined color, say blue; if incorrect, it changes to a different predetermined color, say red. This change of color simultaneously records his response and shows the learner whether his choice of answers was correct or incorrect.

Obviously this device can be used both as a means of guidance in learning and as a means of improving speed and accuracy in scoring tests of what has already been learned. When used for guidance, questions are made more minute and searching and the learner is permitted, after reading a question, to make the necessary calculations and investigations—which may include the reading of reference or text material—before choosing and recording his answer. This was the manner in which the device was used in the present study.

The reading matter here used consisted of Chapter VI and a portion of Chapter VIII of Woodworth's *Psychology* (revised). This reading matter was divided into three separate assignments as follows: (a) Assignment A comprising the first half of Chapter VI accompanied by 42 appropriate questions, (b) Assignment B comprising the remaining half of Chapter VI accompanied by 42 questions, and (c) Assignment C comprising the first 21 pages of Chapter VIII accompanied by 25 questions.

In Assignments A and B the subjects were all members of two sections of a class in introductory psychology, who were at liberty for two consecutive hours between 1 and 5 o'clock on certain afternoons in November 1930. In Assignment C the subjects were all those who took part in previous assignments and were present at a given class period later.

The subjects were divided into an experimental and a control group of equal ability as rated by previous tests in psychology. In Assignment B the experimental and control groups of the previous assignment exchanged places so that the controls of the first assignment became the experimental subjects of the second and vice versa. Though the parity of the groups was slightly disturbed by the absence of some subjects, no subsequent eliminations were made to restore exact equality because of the small number of subjects included. That the disparity was not serious in any case will be seen in the record which follows.

Before the commencement of study on Assignment A, and likewise on Assignment B, its accompanying list of multiple-choice questions was taken by all subjects as a preliminary test. After leisurely completing this test, each subject exchanged it for another copy of the same list of questions which he then proceeded to answer to the best of his ability with the aid of his textbook. Students were encouraged to read with care and to re-read as often as necessary. With the understanding that quality alone would determine the score, each student recorded the time of starting and that of finishing the assignment to the nearest five seconds by reference to a large laboratory clock. As soon as he had finished the assignment, each student exchanged his question sheet for a third copy of the same test which, as in the preliminary test, he now answered without reference to his textbook or his notes.

The assignments given to paired experimental and control groups were identical in all respects except that the second test sheet (study sheet) of the experimental subjects was treated chemically as described above, and answers were recorded by

---

This article, written by Dr. Peterson at the Kansas State College, Manhattan, Kansas, is reprinted, with minor editorial changes, from *Transactions of the Kansas Academy of Science*, Vol. 34, 1931.

[1] Devised by H. J. Peterson and J. C. Peterson and previously described by the latter as "A New Device for Use in Teaching, Testing and Research in Learning" in the *Transactions of the Kansas Academy of Science*, Vol. 33, 1930, p. 41-47. (A patent on the device is pending.)

means of a "chemopen"[2] instead of the ordinary pen or pencil used by the control subjects. The subjects were encouraged to do their best in all tests and, just preceding the second and third test on each assignment, were informed that scores in these tests would count in the determination of their semester grades.

In the presentation of results the following abbreviated terms are used as defined below:

M—The arithmetic mean
C—The control group
E—The experimental group
I—The preliminary test
II—The same test answered during study with the aid of text
III—The same test repeated immediately after study without reference to the text.

Table 1. SCORES ATTAINED IN ASSIGNMENT A[3]

	GROUP E (EXPERIMENTAL, N=23)		GROUP C (CONTROL, N=20)	
Test	M	sM	M	sM
I	27.00	.63, r_{13}=.39	26.85	.93, r_{13}=.69
II	33.00		30.25	
III	37.61	.90	30.40	.90

Disregarding for the present the scores attained in Test II during study with the text, let us observe the gains made by Group E and Group C in Test III as compared with Test I. Group E gained 10.61 points and Group C gained 3.55 points. Those who used the self-checking device thus gained three times as much, on the average, as those who used only the test questions for guidance in reading. The difference in gains is 7.06 points in favor of the experimental group, and the standard error of this difference is 1.21 points. The difference in the gains of Group E and Group C is therefore 5.83 times as large as its own standard error in favor of the experimental group.

---

[2] The chemopen is a strip of absorbent material (here felt) moistened with water or, when needed, a chemical in solution and enclosed for convenience in a watertight tube.

[3] Throughout this paper the symbol "sM" designates the standard error of a mean.

In Assignment B, as previously stated, the groups exchanged places so that the experimental group of Assignment A became the control group of Assignment B and vice versa. The results of this experiment are listed below in Table 2.

Table 2. SCORES ATTAINED IN ASSIGNMENT B

	GROUP E (N=20)		GROUP C (N=22)	
Test	M	sM	M	sM
I	27.18	.62, r_{13} = .30	26.95	.72, r_{13} = .68
II	31.86		30.90	
III	37.91	.82	31.41	1.00

In this assignment Group E gained 10.73 points, as compared with a gain of 4.46 points for Group C in passing from Test I to III. Here, then, the group that gained only one-third as much as its competitors through study of Assignment A gained 2.4 times as much as those same competitors through study of Assignment B. This marked shift in gains always in favor of the group who used the self-checking device when other factors were constant, must apparently be attributed to the influence of the device on learning. The difference between the gains of the experimental group and the control group is here 6.27 points, and the standard error of the difference is 1.16 points. (The reliability of the difference between gains is here calculated by means of the formula

$$sDg=\sqrt{s^2_1+s^2_2+s^2_3+s^2_4-2r_{12}s_1s_2-2r_{34}s_3s_4}$$

from Lindquist, E. F., and Foster, R. R., *Journal of Educational Psychology*, Vol. 20, 1929, p. 105.) This difference in gains is therefore 5.40 times its own standard error.

Fortunately, similar comparisons can be made between the performances of each group with and without the use of the tester because Assignment A and B were found to be almost exactly equal in difficulty. In the preliminary tests on these assignments the mean score of the smaller group was one-third of a point higher in B than in A, while the mean score of the larger group was one-twentieth of a point lower in B than in A. Comparing the two mean scores of the larger group on these preliminary tests, we find an advantage of 6.71 points in the test with which the self-checking device was used. The standard error of this difference is 1.48 points. The difference is therefore 4.53 times its own standard error. Comparing the two mean scores of the smaller group on the preliminary tests, we find an advantage of 6.05 points in the test with which the self-checking device was used. The standard error of this difference is therefore 4.91 times as large as its own standard error.

Notwithstanding the smallness of our groups, all of the four comparisons of performance listed above favor the group using the tester device or the test with which the device was used. And each group and each test was in its turn so favored. The smallest of these four differences is four and a half times as large as its own standard error. This gives a fairly strong presumption of genuine and probably large value in the use of the device as a guide in reading for information. But the fact that the final test was a mere repetition of the preliminary and study tests may raise a question concerning the flexibility and applicability of knowledge gained by its use. A partial test of this point was made in Assignment C.

The purpose of Assignment C was to test the flexibility of knowledge gained as a result of using the tester device. The final test here was accordingly formulated in new terms and the majority of its questions stated in completion-test form. To introduce still further variation, its questions were presented orally. The preliminary test was omitted and the study test reduced to 30 minutes duration. To motivate the home preparation, students were told two days prior to the test, when the assignment was announced, that there would be an objective test before class discussion. Here, as in Assignments A and B, students were allowed to use their texts while taking the study test but not during the final test.

The results of this experiment are presented in Table 3 below.

Table 3. SCORES ATTAINED IN ASSIGNMENT C

	GROUP E (N=17)		GROUP C (N=19)	
Test	M	sM	M	sM
II (Study test)	17.53	.85, $r_{23} = .90$	19.05	.51, $r_{23} = .71$
III (Final test)	20.59	1.17	18.84	.75

Apparently the difference in the mean scores of Group E and Group C in Test II results from the fact that more members of the former than of the latter group failed to finish the test before time was called. Thirty-three questions remained unanswered on the test sheets of six subjects in Group E, whereas all members of Group C finished except one who omitted five questions.[4] The gains from Test II to Test III were +3.06 for the experimental group and -.21 for the control group. The difference in these gains is 3.27 in favor of the experimental group, and the standard error of this difference is .762. The difference is therefore 4.30 times as large as its own standard error.

Here again the difference in gains in favor of the subjects who used the tester device is statistically significant. Moreover, despite the limitations in time which hampered the experimental group more seriously than the control group, the difference is almost as significant here as in the preceding assignments. Apparently the flexibility of knowledge or its availability for use in the solution of new problems is not adversely affected by the use of our tester device in its acquisition.

The value of the Self-Instructor and Tester is, moreover, not limited to the self-checking feature which alone accounts for the facilitation of learning found in the foregoing experiments. The entire device also comprises numerous objective questions which serve to direct the attention of the learner and motivate his efforts. The value of such questions is roughly indicated by the findings of Washburne (1929), who reported that "a genuine difference in learning amounting, in the case of experimental generalizations, to 40 percent of the mean score or over 5 P. E. diff. is involved in the use of questions." In a study of learning which involved 10 repetitions, Symonds and Chase (1929) likewise find objective tests of considerable value. "The value of test motivation," they conclude, "may be estimated as the equivalent of about five sheer repetitions."

In conclusion, it may be said that of the five comparisons here made between performance with and performance without the self-checking feature of the Self-Instructor and Tester, all comparisons showed statistically valid differences in favor of performance with the self-checking feature. On the average, the groups . . . [that] used this feature of the device in reading gained from 2.4 to three times as much in information as did those who used only the questions as a guide. Gains were practically as great when study-test questions were reworded and changed to completion form as when the same multiple-choice questions were given both in the study test and in the final test. Almost invariably students express a strong preference for the entire device, including the self-checking feature, as compared with the mere list of objective questions.

---

[4] In Assignment A an average of 64 minutes was required by the experimental subjects and 63.8 minutes by the controls. In Assignment B an average of 62.4 minutes was required by the experimental group and 60.9 by the controls.

Results of Use of Machines for Testing and for Drill upon Learning in Educational Psychology

JAMES KENNETH LITTLE

THE EXPERIMENTAL study of problems of college teaching is a development quite largely of the last decade. However, these investigations have already yielded suggestions for the improvement of college instructional procedures which should, if effectively combined, markedly increase the efficiency of college work. A still newer development is the invention of various devices for automatically scoring tests and giving drill. Such devices should facilitate experimentation at the college level. Certain of them seem (as will be pointed out shortly) peculiarly suited for following out certain of the suggestions referred to above for the improvement of college work. The very practical problem of the investigation reported in this paper[1] was to see what might be accomplished, using certain of these testing and drill machines, to follow up some of these suggestions for the improvement of instructional efficiency.

MATERIALS AND SUBJECTS

Previous investigations of college instructional problems had (the writer felt) emphasized the following relevant points: (a) the motivating effect upon the learner of knowledge of standing and progress; (b) the value, both for motivation and for guidance in learning, of informing students specifically and immediately of their errors and their successes in their work; (c) the value in the above (and other) connections of the frequent short test, as contrasted with less frequent longer tests or examinations; (d) the great importance of continuous adjustment to

---

This paper, by Dr. Little, of the University of Wisconsin, originally appeared in the *Journal of Experimental Education*, Vol. 3, 1934. The present version has been condensed slightly for this volume by S. L. Pressey.

[1] The paper summarizes a thesis submitted in partial fulfillment of the requirements for the doctor's degree at The Ohio State University. The writer wishes to express his obligations to S. L. Pressey for his assistance in this investigation.

individual differences not only in capacity but also in error pattern and difficulty; (e) the value, in all of these connections, of a consistent use of the make-up test.

Two pieces of apparatus available to the writer appeared especially useful in facilitating an instructional program developing the above points. The first was a device for the immediate automatic scoring of objective tests *and* tabulation of results by item.[2] This machine made it possible for an instructor to score a 30-question test sheet in less than a second. Immediately after he had finished a test, a student could therefore know his score. And as soon as the last student's paper had been run through the machine, the instructor could glance at the item-tabulation dials and at once discuss with the class those questions on which most mistakes had been made. The second, or drill, machine was a small device, one of which was given each student for use during the drill period. The student indicated his answer to each question by pressing one of five keys (one of two, for true-false questions). If the answer was correct, the drum of the machine turned up the next question; if incorrect, the same question remained before the student and he was thus required to find the correct answer to each question before proceeding to the next. Meanwhile, a counter dial counted the total number of tries made in completing the test. The students were shown how to reset the machines and were expected to go through each test until they did so with no errors.[3]

The investigation was carried on during the winter and spring quarters in the five-hour required quarter course in educational psychology and dealt with the regular work of the course.[4] The course was taught, in sections averaging 30 students, by five different instructors; in all, 14 sections took part in the experiment. The students were mostly freshmen and sophomores.

The course was divided into 14 teaching units; for each, two 30-item true-false tests were available. There were also used: (a) a 140-item pretest (reliability .80) covering the entire course and given to all sections at the beginning of the quarter, (b) a 100-item selective answer (five choice) midterm covering the first half of the course, (c) a similar 100-item test covering the second half, and (d) three broad essay-type questions sampling from the entire course. In working up results, the midterm and final objective examinations were combined and considered one final test (reliability .92).

---

[2] For a description, see Pressey (1932).

[3] For a description of an early form of this machine, see Pressey (1926).

[4] Various preliminary investigations were made, using nonsense syllables, German vocabulary, and unfamiliar Tartar history. They suggested progressively greater value of the procedures used according as the material was meaningful. Because of this and because of the greater practical significance of experimenting with regular subject matter, the major effort of the experiment was therefore directed as above described.

PROCEDURE

A calendar was prepared at the beginning of each quarter. This required that the sections keep together, spending an average of about three days on each teaching unit, and that each test be given on the same day. Twelve such tests were scheduled, the last chapter of the first half of the course being included in the midterm and the last test of the second half in the objective final.

Each quarter, two sections of the course were run as "test-machine groups," two as "drill-machine groups," and the remaining three sections as control groups; there were thus, in all, four test sections, four drill sections, and six control sections. So far as possible, sections of each type were rotated among the instructors and times of day so that differences in instructor efficiency, time of day, etc., would be equalized; and careful examination of the data would indicate that such factors were largely negligible or thus equalized. The pretest and the University Intelligence Test (given to all entering students) made it possible to pair cases as to ability and preparation.

All sections took pretest, midterm, and finals. In the test-machine sections each one of the 12 unit tests was scored by the scoring and tabulating machine; as each student finished punching his answer slip, he brought it to the instructor, who at once scored it, checked wrong answers (very easily done when the slip is in the machine), tabulated the score on a tabulation form already on the board, and returned the slip to the student. When all had finished, the test was discussed, the instructor directing the discussion toward the items which the item-tabulator revealed as most frequently missed. Each student thus at once knew his score. Since all the tests were 30 questions long, he readily sensed how far he was from mastery of the material. From the distribution on the board he could see what his relative standing was; he could also see his letter-grade mark, since these were also indicated on the distribution. Further, all students who on the first unit test scored below B were required to take a make-up test (the duplicate form of the first test) on the following day, and this was scored and discussed in similar fashion. The final mark on the unit was then taken as the average of the first and make-up tests. The ways in which this procedure provided motivation by immediate knowledge of results and progress, informed of errors, and adjusted to individual difference will be evident.

In the drill sections the same general procedure was used. But the tests were taken using the drill machines. The score the first time through was tabulated as the student's score on that test, but the student then reset the machine and repeated the test until he could go through it without error. The drill sections thus had the advantages of the test section, plus review with automatically required corrections of error, and opportunity for immediate progress to mastery of that test.

The control sections took the chapter-unit tests by marking on an answer slip which was graded by the instructor that evening and returned the next day. There were no make-up tests. The control sections thus did not have immediate knowledge of results (they did know results the next day, and the cards had errors checked so that they could locate mistakes). They did not have the stimulus of possible excuse from the make-up if they did well on the first test, the motivation involved in the possibility of improving their mark if they did poorly in the first test, or the help which taking the make-up and locating any errors might give them. Nor did they have the benefit of the diagnostic reviews with measure of progress provided by the drill machine. The control sections had this possible *advantage*, that the extra time taken by the testing and the make-ups in the experimental sections could in the control sections be devoted to teaching and class discussion.

RESULTS

In handling results, each student in an experimental section was paired with a student in a control section on the double basis of score on the pretest and on the intelligence test; pairing was fairly close (within five points) and, in consequence, the following data do not, of course, include all cases. Table 1 shows the combined (winter and spring quarter) results for the "test-machine" group and its control and the "drill-machine" group and its control, on both the 200-item objective final tests and the 45-minute, three-question essay final examination. Not

Table 1. PAIRED CASE RESULTS, SPRING AND WINTER QUARTER COMBINED

200-ITEM MULTIPLE-CHOICE FINAL TEST						
	TEST			DRILL		
	Exp.	Con.		Exp.	Con.	
185-9	1			1		
180-4	1	1		3	1	
175-9	1	2		2	1	
170-4	8	7		4	3	
165-9	8	6		6	4	
160-4	11	3		7	5	
155-9	9	6 --Q_3		9	8 --Q_3	
150-4	10	5		8	5	
145-9	10	13		10	1	
140-4	12	8 --M		6	4	
135-9	6	7		3	4 --M	
130-4	6	11		3	8	
125-9	9	9 --Q_1		2	8	
120-4	4	5		3	6 --Q_1	
115-9	1	4		3	3	
110-4	1	5		1	3	
105-9		4			4	
100-4		3			2	
95-9					1	
90-4						
85-9	1					
N =	99	99	diff	71	71	diff
Q_3 =	163	156	7	163	158	5
M =	150	141	9	153	136	17
Q_1 =	138	127	11	143	124	19
ME — MC	9				17	
PE or diff	1.92				2.34	
Diff/PE	4.69				7.26	

45-MINUTE, THREE-QUESTION, ESSAY-TYPE FINAL EXAMINATION						
	TEST			DRILL		
	Exp.	Con.		Exp.	Con.	
60-2						
57-9				1		
54-6	4	1		1		
51-3	1	2		2	1	
48-0	5	7		4		
45-7	5	4		5	4	
42-4	13	3		11	3	
39-1	7	11 --Q_3		7	6	
36-8	10	11		7	8 --Q_3	
33-5	10	9		2	12	
30-2	11	9 --M		14	9 --M	
27-9	14	10		4	5	
24-6	5	11 --Q_1		4	5 --Q_1	
21-3	4	7		3	5	
18-0	6	5		3	3	
15-7		7		1	7	
12-4		1		2	3	
9-11	1	1				
6-8	3					
3-5						
0-2						
N =	99	99	diff	71	71	diff
Q_3 =	43	40	3	44	38	6
M =	35	32	3	37	32	5
Q_1 =	28	25	3	30	24	6
		3			5	
		1.06			1.17	
		2.83			4.28	

(ME=median of experimental group; MC = median of control group.)

End of quarter differences between (1) test-machine group and its control, and (2) drill-machine group and its control, cases being paired on the basis of (a) intelligence and (b) a pretest on the subject matter of the course.

only the medians but also Q_1, Q_3, and the entire distributions are given, since the differential effect of the experimental procedures upon the lower as compared with the upper part of the distributions is the most striking feature of the results. Thus it will be noticed that on the objective final test, Q_1 of the drill group is 19 points above Q_1 of the control (seven points above the median of the control), whereas the experimental Q_3 is only five points above the corresponding control. On the objective test the drill group shows greater gains than the test group. On the essay-type final this difference also appears, but all differences are less. This is probably due in part to the unreliability of the grading of essay-type questions. At the bottom of the table are shown the differences between experimental and control medians, the probable errors of these differences, and the differences divided by the probable errors. It should be again emphasized, however, that the most important differences are at Q_1 rather than at the medians and that the results, and their reliabilities, are to be evaluated from an inspection of the entire distributions rather than from any one comparison.

. . . . . .

THE LARGER SIGNIFICANCE OF THE STUDY

The investigation here reported was planned as a very practical class experiment. A variety of factors which might presumably contribute to increased efficiency were therefore included, and analysis of the contribution of each is hardly possible. However, two general conclusions of very considerable importance may, the writer feels, be drawn. In the first place, the experiment is (so far as the writer knows) the first instance of the systematic use of machines as part of the regular routine in regular university classes. It demonstrates (he feels) the possibility of mechanical aids in education.

In the second place, the experiment is in interesting contrast to present trends in efforts to improve college instruction. The present trend seems to be toward the expensive tutorial system, toward still more expensive housing plans, and toward somewhat transcendent or sentimental emphasis upon the importance of "personality" in teachers. In contrast, one purpose of this experiment has been to show that an impersonal attempt to organize procedures in college instruction for greater efficiency may produce demonstrable results and, at the same time, save labor. In this time of depression such a saving may be of some interest.

CONCLUSIONS

1. Students immediately apprised of their test results, and given opportunity to correct deficiencies by make-up tests, profit markedly in terms of final examination results over students who do not have such advantage.
2. Students immediately apprised of the correctness or incorrectness of their responses to each item of a test, and given opportunity to correct deficiencies by drill and by make-up tests, likewise so profit.
3. The greatest benefit accrues to students who usually score in the lower half of the distribution, although the entire groups move upward.
4. Mechanical test-scoring and drill devices have a practical use in the classroom. They are convenient for students, time- and laborsaving for teachers, and make possible instructional techniques not otherwise practicable.

A New Self-Scoring Test Device For Improving Instruction

GEORGE W. ANGELL and MAURICE E. TROYER

In 1926 S. L. Pressey reported a mechanical instructor which automatically and immediately gave each student knowledge of the correctness of his response to objective questions. By 1932 Pressey reported the development of two more test-scoring machines. One was a small cash-register-like machine on which a student gained immediate knowledge of results upon his response to each item in a test. The other was a test-scoring machine that enabled the instructor to score and item-analyze short tests as rapidly as students completed them during the test period. J. K. Little's research (1934) utilizing these machines indicated that students who gained immediate knowledge of test results via these methods scored higher on objective and essay final examinations than did students whose papers were scored by the instructors and handed back the following day.

In 1930 J. C. Peterson described two procedures by which students might gain immediate knowledge of results while taking examinations. By the one method, students using a pin punched through an envelope and several layers of cardboard with selector holes in such a way as to record and discover their right and wrong responses. By the second method, students gained immediate knowledge of results by moistening the ink-spot of their choice. Correct answers turned blue, incorrect responses turned red.

During World War II the United States Armed Forces utilized several devices that gave students immediate or early knowledge of test results. One such device was the Automatic Rater, a machine that weighed 175 pounds, was electrically operated, and had over-all dimensions of 18" x 26" x 48". An individual device of this magnitude is an expensive proposition and indicates the importance placed on immediacy of knowledge of test results by the Armed Forces.

The value of such testing procedures for the learner appears obvious in the light of past research. Immediate knowledge of test results is consistent with the leading theories of learning as proposed by E. L. Thorndike (1932), E. R. Guthrie (1935), C. L. Hull (1943), and E. C. Tolman (1932). Yet, traditional practices of test administration will not be abandoned until simple, inexpensive, rapidly prepared, foolproof devices are made available for general use. In an effort to develop such an apparatus, the authors have been experimenting with various types of punchboards since September 1945.

One of the most promising types of punchboards has been experimentally produced and used for research studies during the past two years. This particular instrument has three parts: (a) the front cover with five perforations for each item of the test; (b) a thick cardboard middle section perforated in the same manner; and (c) the back cover which has no perforation. In order to prepare the punchboard for use, an answer sheet (8½" x 11") is inserted between the front cover and middle section. A "key" is then inserted between the back cover and middle section so that red spots for the correct answers to the item will show through the proper perforations of the front cover and middle sections. For some test situations, indicia such as letters or numbers are more appropriate than red spots. In either case, the student punching a hole in the answer sheet with his pencil may look into the socket to see if the keyed indicium appears. In this manner a student may select his answer for Item 1 and punch the corresponding spot on the punchboard. If the key shows that his answer is correct, he may go to Item 2 or any other item he desires. If the selection for Item 1 is wrong, he may look over the remaining responses and continue to punch until the correct answer is located.

A cardboard punchboard is relatively inexpensive. It may be used repeatedly with different keys. Keys may be mimeographed or offset by the hundred. A person can handily carry 50 to 75 loose punchboards or 200-250 of them when packed in an average-sized carton. There is nothing mechanical to go out of adjustment during the examination. Item analyses can be made from the answer sheets. Students may answer the items in any order as well as look back at items to see where

---

This paper is reprinted from *School and Society*, Vol. 67, No. 1727, January 31, 1948.

mistakes were made. For short quizzes, a student, upon completion of the test, may drill himself the remainder of the period by handing his answer sheet to the instructor and sliding in a fresh one.

This punchboard has been used experimentally in chemistry and citizenship classes at Syracuse University during the 1946-47 academic year. Data from both studies showed that learning is significantly enhanced by immediate knowledge of results through the punchboard-testing technique.

A device which allows students to evaluate immediately their responses to test items opens up a number of possibilities for teaching improvement. As far back as 1929 F. D. Curtis and G. G. Woods (1929) reported the superiority of students' grading their own papers. Pressey is currently experimenting with special self-instructional quizzes to accompany various units of study in psychology. Peterson (1931) has shown the value of self-scoring tests to accompany reading assignments. The utilization of self-scoring instruments is entirely consistent with the philosophy underlying the ever-increasing "child-centered" school programs.

In general, techniques whereby a student gains immediate knowledge of test results aid directly in the realization of a major goal of education: to help the student to increased ability in self-evaluation, to identify his own strengths and weaknesses in such a way that he may direct subsequent learning efforts more intelligently.

Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-Instruction

S. L. PRESSEY

THE GRADING of conventional essay-type examinations has long been recognized as burdensome and slow. By contrast, the scoring of objective tests has seemed easy and prompt. But actually this latter task is by no means negligible or as immediate as might be desired. If test materials or simple testing devices could be developed so that *as a student answered each question, that answer was immediately and automatically scored and recorded as right or wrong*, then clearly much trouble would be saved. Additional advantages of self-scoring devices are more important, however. A test in French or mathematics or navigation is of little value if it does not in some way further the student's learning of that subject. If he is weak on certain points, the test should locate them and aid in the remedying of these weaknesses. And this should be done *promptly*. Self-scoring devices *inform the student immediately—when he indicates his answer to an objective question—whether that answer is correct; if it is wrong, he is at once automatically guided to the discovery of the correct answer.*

As an instructional device, the self-scoring test has certain further advantages. An objective test can cover many more questions in a given time than a written examination or quiz. The self-instructional objective test can similarly cover more ground than would be possible otherwise. It can in effect give individualized instructions to every student in a class (searching out the misconceptions of each student and then correcting

---

The original version of this paper was published in the *Journal of Psychology*, Vol. 29, April 1950. The present version has been edited and condensed by the author. The work described in this paper was under sponsorship of the Special Devices Center of the Office of Naval Research, in cooperation with The Ohio State University Research Foundation, under the direction of the author.

them) over more questions than a tutor could cover with one student. A well-made objective test has a more carefully considered and exactly phrased right answer than an instructor is likely to give in class discussion, and the wrong answers are carefully chosen as the most common misconceptions or mistakes. The self-instructional procedure systematically presents these for consideration and then clearly checks them off as wrong. In short, these devices should make it possible to cover more ground in a given time than a teacher could *and* do it more adequately.

The purpose of the present paper is, briefly, (a) to describe means by which tests may be made self-scoring and self-instructional, (b) to investigate the way in which such devices work, and (c) to determine their value in educational or training programs.

MATERIALS AND METHODS

The total project has involved development or improvement of several different devices. However, the data of the present paper were all obtained with the "punchboard." This is a very simple unit, the size of a 3" x 5" card, about one-half inch thick. A center of quarter-inch plyboard has riveted, on either side, two thin sheets of pressboard. Between the outside sheet and the second sheet there is sufficient space for a slip of paper to be inserted; between the second sheet of the pressboard and the plyboard center is a removable key sheet of pressboard. On each face of the punchboard are two columns each of four rows of one-eighth inch holes, the rows of four holes being numbered from 1 to 30. However, the key sheet has holes only for the right answer on the particular test to be used. The accompanying plate shows the device with a key sheet and used slip (Figure I).

Test questions are usually presented on a mimeographed sheet. The testee takes a test by simply punching with a pencil point through the paper slip in that hole which corresponds to what he thinks is the right answer. For instance, if for Question 1 of the test he thinks the second answer is right, he will punch his pencil through the paper appearing in the second hole of the first row. If he *is* right, his pencil goes through the paper and down into the hole in the key sheet. But if he is wrong, the pencil barely breaks the paper and then comes up against the key sheet. He thus knows that he is wrong and tries another hole in Row 1, thus proceeding until the pencil goes deep.

After he has gone through a test, he can glance back and

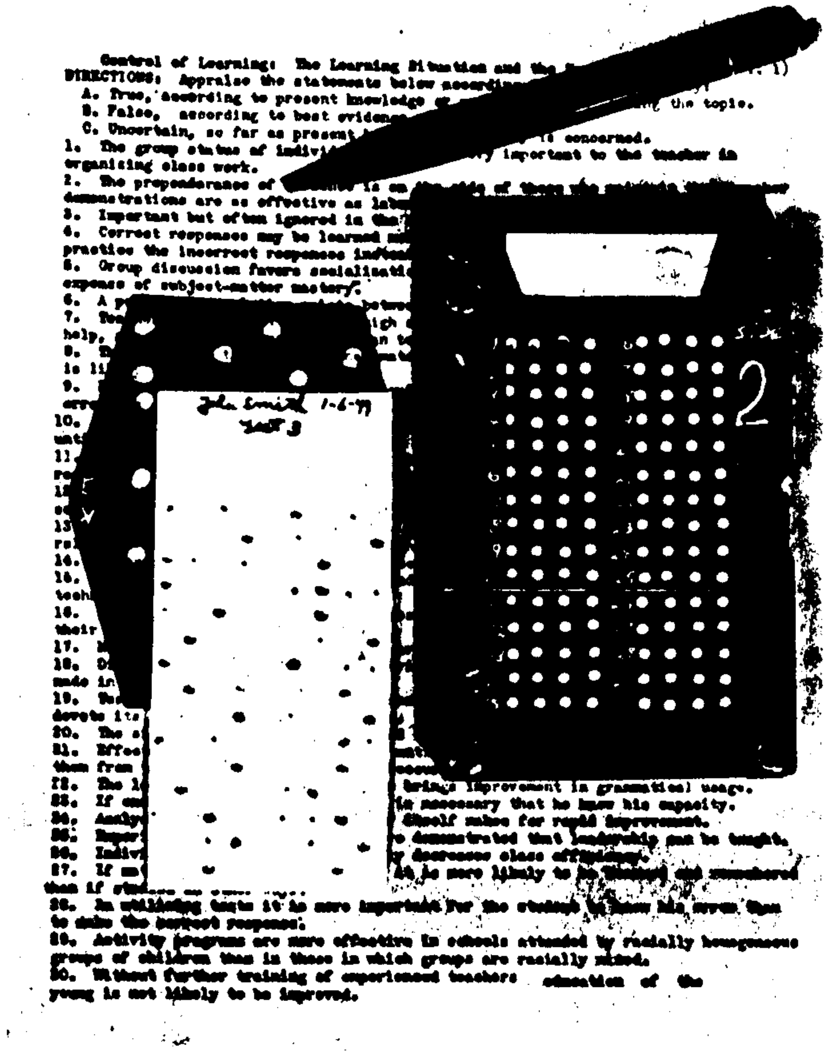


FIG. I. Punchboard with punched paper slip and key withdrawn from the punchboard after use with a mimeographed test sheet.

see on what questions he made a mistake (where he punched more than one hole or has a small hole). The punched papers, once withdrawn, can be filed and are a convenient record for item analysis or other study. The keys are so planned that each can be turned over or turned end for end and thus give four different patterns. It is relatively easy to punch a new key if one is desired. For true-false or other two-choice tests, four columns can be used on one side. The back side of the punchboard, exactly like the front side but numbered "2," either permits a second trial on the test to find improvement and give more practice or provides a total of 60 multiple-choice items.

The tests used with the punchboard in the experimentation to be reported almost all consisted of four-choice objective questions. With a little careful planning, a test of 30 four-choice questions could be put on one side of a single long mimeographed sheet, or 30 true-false statements on one side of a short sheet, and almost any test of that many items on two sheets or two sides of one. In addition, the punchboard could be kept small. For use on the arm of the usual classroom chair, such compact convenience was a distinct asset.

Initial experimentation was with a great variety of subject-matter. It was finally decided to concentrate on material of three types. Russian vocabulary was used as rote matter more interesting and valuable than nonsense syllables, equally unfamiliar to almost everyone, and indicative of possible values of self-instructional tests in learning a foreign vocabulary. Hard English vocabulary items were used to represent tasks such as the learning of technical terms where the material was highly meaningful but without any organization or structure. And the subject matter of two psychology courses served as an example of material which was highly organized and part of a systematic program of instruction. For some of these tests, several forms were available, using different keys. For the two courses, sufficient variety in material was sought to give interest and also to prevent a routine approach. Thus for one topic there were two four-choice tests, a true-false–uncertain-judgment test, and an application test consisting of a series of paragraphs, each presenting a practical situation, or episode, followed by questions as to ways of best dealing with such a situation. There were also various check tests to determine the effects of the practice tests, which will be described later as used.

At The Ohio State University the first and second courses in elementary general psychology and the first course in educational psychology are taught in sections of about 35 students, mostly freshmen and sophomores. In a given quarter there may be a total of as many as 50 sections. Most of these are taught by part-time instructors in their third or fourth year of work for the doctorate. Each course is under the supervision of a senior staff member; there is a regular calendar, certain objective examinations are taken by all sections, and there are frequent staff meetings. These numerous sections, thus coordinated and systematically appraised, offered unusual opportunity for the extensive group experimentation desirable in this project. Further, it was possible in the course in educational psychology to make sweeping changes in methods and hours of meeting when needed for experimental purposes. In short, for the three quarters of the regular school year, the total number of cases, sections, and facilities for this investigation seemed exceptional.

RESULTS

In attempting to appraise the value of devices involving immediate automatic scoring and self-instruction, the natural first step was to determine whether, in fact, taking a test by such means would at once bring about learning. A natural next step was to investigate whether such devices were of service in regular instructional or training programs. A desirable third step was to see whether self-scoring tests might be exceptionally useful with special groups or instructional plans. These issues will be taken up in order.

1. Immediate Outcomes of Use of the Punchboard

As mentioned earlier, Russian vocabulary, hard English vocabulary, and subject matter in psychology (as illustrative of respectively, rote matter, meaningful but unorganized material, and material organized and meaningful) were chosen to show how the punchboard worked with matter of various types. Certain immediate outcomes of such use are now to be considered.

EFFECT OF USE OF THE PUNCHBOARD ON NUMBER OF ERRORS WHEN A TEST IS IMMEDIATELY REPEATED. If students take a test in the usual way and then at once take it again, they would be expected to make about as many errors the second time as the first. In fact, sometimes they do a little worse the second time, as a result, perhaps, of boredom and carelessness. When a class took a 30-word Russian vocabulary test in the usual way (the students simply checking what they thought the right answer) and then were immediately given the same test again in the same way, the median number of errors increased from seven to eight. But a class which took this same test twice using the punchboard showed a drop from 12 to eight errors. With meaningful material, the learning with the punchboard was even greater: from 43 to seven errors on the English vocabulary test and 18 to two on the psychology quiz.

It seems clear that the students do show some immediate learning as a result of using the punchboard. Apparently they gain more with meaningful material and least on matter of a rote type such as the Russian. The great gain in the English vocabulary might be explained as due to the great initial difficulty of these terms plus their basic familiarity, making them easily learned once the punchboard gave reminders as to their meaning.[1]

EFFECTS OF THE PUNCHBOARD AS EVIDENCED BY A TEST DIFFERENT IN TYPE FROM THE PRACTICE TESTS. Conceivably, learning brought about by the punchboard might be highly specific—perhaps a mere memorizing of the test material or the key. If this were true, a second test different from the practice material would show little gain. The following simple experiment offers evidence that the learning was not so limited.

The multiple-choice Russian test used previously was first given a class, using the punchboard. Then a very different test, a recall test, was given. On this recall test, the Russian words were listed in a different order from that in which they appeared on the practice test, and a space was provided after each Russian word for the student to write its English meaning. This same recall test was given also to another similar class after the practice test had been taken simply as a test. Check tests different from the practice tests were also devised in English vocabulary and psychology, and similarly used.

Effects of the punchboard are here not so great or so clear-cut. A repetition of exactly the same test would inevitably show more gain than check tests which vary from the instructional test. However, gains still appear on the Russian vocabulary, and on the meaningful English vocabulary and psychology they are substantial.

EFFECTS OF SEVERAL CONSECUTIVE REPETITIONS OF A TEST USING THE PUNCHBOARD. If once through a practice test with the punchboard does bring some learning, several times through should bring more. However, if students go through the same self-instructional test several times in succession, it is conceivable that they might tend to learn the key or pattern of right answers rather than the information conveyed in the test. To check on these points, one class was given the same Russian vocabulary test practice test four times using the punchboard and then a recall check test, all in one class hour. A similar class was then given a series of practice tests covering the same words but with the order of items and the key or pattern of right answers changed each time; again the punchboards were used but changed each time to fit the key for that particular test. And again at the end of the series the recall end test was given. Similar series were run with same and changed tests in English vocabulary and psychology, except that the length of the hour permitted only three times through the practice series and left no time for an end test in psychology. Table 1 shows the median

Table 1. DECREASE IN NUMBER OF ERRORS WITH REPEATED TAKING OF A PRACTICE TEST USING THE PUNCHBOARD, AND EFFECTS OF CHANGING THE ORDER OF QUESTIONS AND THE PATTERN OF RIGHT ANSWERS UPON SUCH DECREASE*

	RUSSIAN VOCABULARY						ENGLISH				PSYCHOLOGY		
	Practice test					Check	Practice test			Check	Practice test		
	1	2	3	4	5	test	1	2	3	test	1	2	3
Median no. of errors													
Same test	11	8	5	3		10	42	14	5	6	17	4	1
Changed test	13	15	9	9	6	10	41	17	8	5	18	3	1

*Numbers of cases for Russian groups were 49 and 28, for English 29 and 41, and for psychology 33 and 34. Error scores on end tests are mistakes plus omissions. Five minutes' study of the Russian vocabulary list preceded the practice tests.

---

[1] Detailed analysis of the above and other similar results indicated that the punchboard brought improvement in two ways. When a right answer was at once hit upon the first time through a test, the punchboard confirmed this choice as right. But this effect was not very important. It brought gains over what occurred simply with repetition of the test as a test of only about 5 percent. The big effect of the punchboard was in correction of errors. The punchboard brought an immediate right answer on the second trial for about 65 percent of items wrong on the first testing.

number of errors on each practice test and in the Russian check test.

On the Russian, the changing practice tests seem actually to handicap the learning, in that the median number of errors actually goes up on the second trial and thereafter continues higher than on the tests in which the order was unchanged. However, the check test shows no difference at the end of the practice series. On . . . English vocabulary and psychology, results for same and changed practice material are practically the same.

A natural interpretation of the above results is that with such rote learning, as in the case of the Russian, there tends to be some dependence on the arrangement of items. However, in proportion to the meaningfulness and integration of the material, change or sameness of key makes no difference; attention is on the subject matter, and that is what is learned. Apparently with none of these subject matters does repetition of the practice test with the same key handicap learning. It is even conceivable that the same test and key in successive reviews might help. For example, in reviewing a book it is an aid that one can each time find a topic on the same page with same context, and upsetting if a review is of a different edition where all these details are different. The analogy is, of course, only very rough, but it does seem suggestive of value in review of the same test. Some value would seem especially likely if the test questions were in a logical order and perhaps grouped under headings. Anyhow, these findings are indeed pleasant, since they support the more convenient repetition of the same test for practice.

2. The Punchboard in a College Course as a Testing Aid Not Integrated with Teaching

The most simple, but obviously inadequate, way to use a device like the punchboard is as a convenience in testing but without any attempt to tie its instructional potentialities in with teaching methods. Such a situation might occur if its use were required of teachers without interesting them in it or indicating ways of using it to aid instruction.

In the large second-quarter course in general psychology, eight chapters of the textbook were covered between the first and second midterm examinations. To help students in reviewing for the second examination, two 60-item review tests were constructed. The first consisted of 20 multiple-choice questions on the first two chapters, 20 on the second two, and 20 more which were the first 20 repeated but in a different order. This review test was given to four sections totaling 114 students during the class hour preceding the midterm examination. A second review test similarly covering the second four chapters was given to six other sections, totaling 182 students. On the day following, all sections were given an objective midterm examination, having 21 questions repeated or paraphrased from the first review test, 22 similarly from the second practice test, and 24 questions which were on the same eight chapters but unlike questions on either review test. Table 2 indicates that use of the punchboard reduced errors most on the repeated questions. But some gain appeared also on those which had been rephrased from the practice test to the midterm.[2]

Table 2. PERCENTS OF STUDENTS GETTING QUESTIONS RIGHT ON MIDTERM WHICH THEY (A) HAD TAKEN ON A PRACTICE TEST USING THE PUNCHBOARD AND (B) HAD NOT TAKEN ON A PRACTICE TEST

	NOT TAKEN IN PRACTICE TEST	TAKEN WITH PUNCHBOARD	GAIN FROM PUNCHBOARD TEST
Questions paraphrased from practice test to midterm	69	72	3
Questions repeated from practice test to midterm	69	87	18

*The table shows that 72 percent of the students got questions right on the midterm which they had had in different wording on a practice test using the punchboard. However, only 69 percent of students got these questions right if they had not taken them in the practice test. The punchboard practice test thus brought a gain of 3 percent; in repeated questions the gain was 18 percent.

3. Use of Practice Tests with the Punchboard as a Major Instructional Method by the Instructor in Charge of the Class

This experimentation was in the course in educational psychology—which follows the course in general psychology—is most commonly taken in the freshman year, is taught in sections of about 35 students, and meets five times a week for one class hour. For this work, a large number of practice tests had been carefully prepared, with special emphasis on problems of judgment and application.

---

[2] The original article contained much further experimentation not included in this much abridged form.

The plan was simple and straightforward. In the spring quarter there were 13 sections of educational psychology. Four of these used practice tests with the punchboard in connection with certain units of the course and were taught by the two instructors who were accustomed to the use of the punchboard. During the first four weeks of the quarter, five units or topics of the course were covered. The punchboard groups began consideration of two of these with a punchboard practice test which was used as a point of departure for discussion. They then took a second practice test on the second day in the same fashion. Two sections took the same practice tests on these units; but instead of using the punchboards, they marked their answers in the usual fashion on answer sheets. As they finished, they discussed the questions among themselves and with the instructor, thus arriving at knowledge of the right answers and clarifying doubtful points. These tests were then collected and the answer slips graded and returned a day or so later. This method probably leaned over backward in an effort to give full opportunity for advantages of the practice tests without the punchboard to be obtained; rather more time was given to test-taking plus discussion of the tests than with punchboard group, and return of the answer slips was a second reminder of the test and opportunity for discussion of them. Seven sections did not use these practice tests in any way, the time going instead for informal discussion. Initial status was appraised on the basis of an aggregate index taking account of ability as tested at entrance to the University, scholarship to date, and grade on the preceding course in general psychology. The three groups were found initially so similar that no allowances for initial differences were needed.

A 75-question multiple-choice midterm was then given all sections. Of these questions, 25 were repeated or paraphrased from the practice tests, 10 were new questions but on the same two units as dealt with in the practice tests, and 40 were questions on the three units not having practice tests. In similar fashion during the second four weeks of the course, four units were considered, on two of which two practice tests were given to the punchboard and test groups. The second midterm was then given. It had 24 items repeated or paraphrased from the practice tests, 16 new questions on the same units, and the remaining 35 on the two topics which were new to all groups.

The bars in Figure II show the scores made by the middle half of the students in the three groups on these two midterms. The total scores from such varying types of questions are obviously the undifferentiated sum of various possible effects of the practice tests. But they seemed worth considering as evidence of outcomes in what might well become a common situation: practice tests used for some but not all the material in a course and examinations which appraise students' work as a whole with-

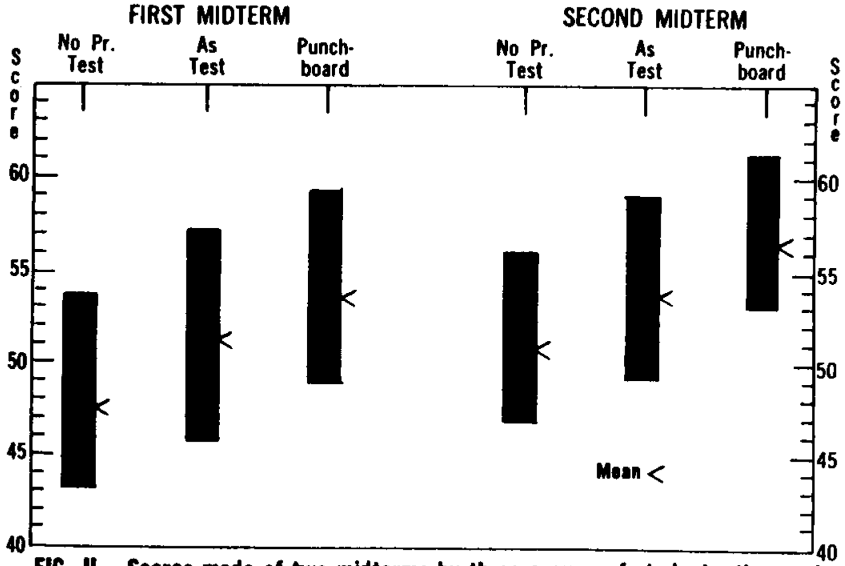


FIG. II. Scores made of two midterms by three groups of students, those who had taken (a) no practice tests, (b) practice tests simply as tests, and (c) practice tests using the punchboard.

There were 225 students in the seven sections not taking the practice tests, 70 in the two sections taking them simply as tests but with discussion after, and 132 in the four sections taking the practice tests with the punchboard. The bars show scores made by the middle half of the students in each group (from the 25 to the 75 percentiles). The practice tests covered only about half of all the topics included in the midterms.

out differentiating portions so aided. The differences between the no-practice-test group and the punchboard groups are evident and clearly significant. Superiority of the punchboard sections over the sections taking the practice test without the punchboard is evident, especially on the second midterm. As always, individual differences in all groups are great.

Analyses of results are shown in Table 3 for all the questions dealing with the units covered by the practice tests and 10 items in each midterm on topics not so covered. The first three columns of figures show average percents for each group passing each type of item on the midterm; the fourth column shows the gain of the punchboard group over the nontest control; and the last column, the gain of the punchboard sections over the "test control" sections.

Table 3. PERCENTAGES OF STUDENTS WHO HAD TAKEN (A) NO PRACTICE TESTS, (B) PRACTICE TESTS SIMPLY AS TESTS, AND (C) PRACTICE TESTS WITH THE PUNCHBOARD* WHO ANSWERED CORRECTLY VARIOUS TYPES OF MIDTERM QUESTIONS

Relation of midterm to practice tests	PRACTICE TESTS			PUNCHBOARD GAINS OVER	
	Not taken	As test	Punchboard	No test	As test
First midterm		Percent correct			
10 Questions not in the practice tests in any form and on different topics	63	61	68	5	7
10 Questions not in practice tests in any form but on same large topics	52	60	62	10	2
20 Items paraphrased from practice tests	66	74	77	11	3
5 Items repeated from practice tests	77	90	91	14	1
Second midterm		Percent correct			
10 Questions not in practice tests in any form and on different topics	78	80	83	5	3
16 Questions not in practice tests in any form but on same topics	65	72	78	13	6
14 Questions paraphrased from practice tests	68	74	76	8	2
10 Questions repeated from practice tests	73	84	89	16	5

*There were 225 students in the seven sections not taking the practice tests, 70 in the two sections taking them simply as tests but with discussion after, and 132 in the four sections taking the practice tests with the punchboard.

The fourth column shows substantial gains of the punchboard group over the group which took no practice tests. The punchboard sections show superiority on new questions dealing with the same units and even on questions on the units not covered by the practice tests. Apparently the discussion and teaching tied in with the tests and aiming at generalization and application had so operated. Moreover, the right-hand, or fifth, column, shows some superiority of the punchboard over the test-control group; and this extends also on both midterms to questions on topics not covered by the practice tests. To what extent these last differences may be considered significant is not clear, but their appearance in all categories surely warrants added confidence in them. And again it must be remembered that the "as-test" control group method . . . gave more than comparable opportunity for gains from tests without the punchboard—more time being given to the tests, plus discussion of them, when used without the punchboard than when used with it. Punchboard testing was also more convenient, time-saving, and interesting to the students than testing in conventional fashion with discussion the next day.

Indications of students' attitudes toward the punchboard were sought. Their liking for it was well evidenced by their protest when, after its use part of a quarter, the experimental design called for its discontinuance. Comments from students were invited. The following samples indicate something as to attitudes and methods of using.

> Every time I punch a wrong answer, I read the question over again. The punchboard makes me think.
>
> When I have the right answer, but am not certain, the punchboard assures me; then I can settle on that answer and remember it, and consider the reasons why it is right.
>
> My errors are pointed out immediately; I'm not left with the wrong ideas.
>
> It's like a game and a challenge; I'm motivated more.
>
> If you punch a wrong hole, you're the only one who knows it; it's not upsetting like a mistake in class discussion.
>
> After I've been through a punchboard test, I feel much more confident, know where I stand, have been checked up.
>
> I find it much easier to ask for help, since the punchboard brings issues to a head.

4. Use of Punchboard Practice Tests with Special Groups of Superior Students

It might be assumed that a device facilitating immediate understanding of test material and self-instruction would be most effectively used by the ablest students. Since a class, like a fleet, tends to travel at the rate of the slower members, any scheme which releases the abler from that restraint should be valuable. Three methods of using the punchboard tests with groups of superior students were therefore tried.

THE ACCELERATE SEMINAR. An accelerate seminar was tried four different quarters. Selection of students for the seminar was based on a brief student data sheet filled out the first day of the course in educational psychology. This sheet covered age, academic record, veteran status, part-time employment, and interest in the seminar as briefly described to the group. Those students who were interested, whose academic records and scores on the entrance test of general ability were high, who were not overburdened with part-time employment, and who seemed to have a reasonable participation in student social life and activities, were then interviewed; and the most promising were chosen for this special group. Average number in the seminars was 26, or a bit less than a tenth of the total 300 or so students taking the course each quarter. Average percentile of the seminar group on the test of ability at entrance was 77 and academic average slightly above B. A majority, but not all, were veterans.

The seminar met only two hours a week (usually in the evening) instead of the usual five one-hour sessions. All the regular readings and projects were covered and, in addition, about a third more reading to compensate for less class time and assure against possible criticism that the seminars were too easy. Three or four punchboard tests on every topic—most with page references to the reading after each question to aid systematic check back on weak points—were made available to the students in a file, and they were shown how to load the punchboards with paper slip and key so that they could do everything themselves. They could thus use the tests for self-appraisal and guidance in small-group discussion at certain class times and also at other hours in the afternoon when the file was open to them. Part of the class seminar hour was given to special projects or other undertakings and to informal discussion at a mature "seminar" level. And since class time was saved, each student was supposed to take an extra course, work part-time, or make other good use of less demand for class attendance. Median course load was 18 hours; and median hours of part-time employment, 10.

To determine the outcome, students were given essay examinations on the extra reading, were graded by the instructor on projects and participation in class discussion, and took the two objective midterms and final given also to the regular sections. The first rows of Table 4 summarize the results in terms of grades on these objective tests made by the regular sections; the accelerate seminar students; and cases in the regular sections paired with the seminar cases as to age, sex, point-hour-ratio through the previous quarter, and grade in the previous course in general psychology. With all these categories, pairing was a bit rough but was reasonably satisfactory.

Table 4. GRADES MADE BY SUPERIOR STUDENTS USING SPECIAL INSTRUCTIONAL METHODS, INCLUDING PUNCHBOARD TESTS, AS COMPARED WITH GRADES OF STUDENTS USING CONVENTIONAL METHODS OF INSTRUCTION

	PERCENTAGE OF STUDENTS OBTAINING A and B Grades	E*	NUMBER OF STUDENTS
27 Regular sections	28	11	801
4 Accelerate seminars	65	2	106
Cases paired with seminar cases	55	5	106
2 Examination-credit groups	96	0	24
Cases paired with above cases	83	0	24
2 Self-instructional laboratories	67	0	24

*A and B grades are high, but E is failing. Grades here given were based on two objective midterm tests plus an objective final examination including a case study test.

High (A and B) grades in seminar were over twice as numerous proportionally as in the regular sections and failing (E) grades rare. Moreover, the seminar students were better than the paired controls who had over twice as much class work. There were more intangible gains; many seminar students spoke of increased capacity for independent work and closer relations with other students and the instructor because of the cooperative informality. The instructor found the seminars very stimulating. It was concluded that two seminars might be taught as easily as one regular section, making a net gain in number of students handled. Obviously, many factors were involved. But it seemed clear that superior and well-motivated students, taught in mature fashion, and with such aids to self-appraisal and review as the punchboard tests, could save class time and do more than the usual total of college work. Moreover, the demands of this course did not handicap elsewhere; a follow-up of these students and their control groups about a year later showed the seminar group slightly more superior in academic record than before. And as a result of an extra course at the time of the seminar, plus extra credit obtained at some other time, a considerable number of these students graduated a quarter or more sooner than would have been possible otherwise.

THE EXAMINATION CREDIT GROUP. Might superior and well-motivated students be put yet more on their own, with yet more guidance and help from the self-instructional tests? Two "examination credit" groups gave an affirmative answer to this question. University rules provide that a student may obtain examination credit for a course without taking it, if he obtains a grade of B or better on an examination covering the course. The question was . . . whether superior students could, if given a little guidance in independent study plus the help of a series of punchboard tests covering the course, pass the educational psychology without any regular class work.

For this little investigation, records to date of students in the prerequisite general elementary psychology were run over before the middle of the quarter to find students doing exceptionally well in that course and otherwise superior. The best of those thus discovered were interviewed to verify their superiority, explain the plan, and see if they were interested. Only those strongly interested and clearly superior were finally included. Average percentile on the entrance test of general ability was 92, average grade to date was high B (a PHR of 3.35), and average age, 20.5; preference was given to veterans. One quarter, nine students were chosen, and another, 15.

As soon as each group was selected . . . [around midquarter], the first of a series of four two-hour meetings was held. In these, procedure was further explained; assignments made in large blocks, each covering about a third of the course; certain laboratory projects briefly demonstrated; and two objective midterm tests and a final given. There was time for a very little orienting discussion. In addition to the four meetings, the total series of punchboard tests was made available during about four hours each week with at least two hours' attendance required. An additional 20 hours each week were designated as office hours when an instructor or assistant was available for consultation and punchboard tests were also available. Work was thus largely independent study but with frequent punchboard test checks and opportunity for individual consultation. Often two or more students were in for the check tests at the same time and worked together.

In the row labeled "examination credit group" of Table 4 results are shown in terms of objective midterms and finals given regular sections also. The following row gives results for cases paired with those in this group in respect to ability and previous work, in the same fashion as for the accelerate seminars. Again, results are excellent.

THE SELF-INSTRUCTIONAL LABORATORY. A third venture attempted even greater flexibility. At the beginning of two different quarters, small groups of very superior and mature students were given the opportunity to take the course in educational psychology in a "self-instructional laboratory." Each student was given a list of all the work in the course, including the usual informal laboratory exercises. Again, a third more reading was asked of the special groups. Five afternoons a week, a special laboratory room was made available to these students, in which all the materials and readings for the course and a complete series of punchboard tests covering the course were provided. In charge was an assistant who was an experienced teacher and thoroughly familiar with the course and the practice tests. The students were told that there would be no regular class meetings. They could come into the laboratory whenever they wished, work with others or not, consult when they wished with the assistant. Each could progress at his own rate, test his understanding of each topic as he finished work on it, easily look up any item on which he had trouble (using the page references after each question), and then check himself again with a different punchboard test. When each student felt ready for a midterm or the final, he could ask for it. Each could thus finish the course at his own rate; all finished by the middle of the quarter. Results are shown in the last row of the table above. Over two-thirds of these superior and mature students made A or B grades. Some used the time saved in part-time employment, others in taking extra courses or in yet other ways. Several reported gains in effectiveness in independent work.

MAJOR CONCLUSIONS

The preceding pages have reported a variety of projects. What generalizations from them seem most important? Four major conclusions appear to stand out from the total investigation.

1. *It has demonstrated a simple way to telescope into one single, simultaneous process the taking of a test, the scoring of it, the informing of students as to their errors, and their guidance to the finding of the right answers.*
2. *The investigation has shown that such a telescoped testing process, which informs each student immediately as he answers each question whether his answer is correct and which guides him to the right answer when he is wrong, does indeed transform test-taking into a form of systematically directed self-instruction.* Repetition of a self-instructional test brought marked reduction in number of errors made. These effects showed also on types of tests different from the instructional test. The learning was primarily of the subject matter and not of the test key. However, use of the self-scoring device in a college course as a somewhat incidental feature in occasional brief tests brought only limited gains in scores on later objective course examinations; and these gains were confined largely to specific items covered in the practice tests. As with other teaching aids, self-instructional tests must be made a carefully planned part of the total instructional program if their values are to be adequately realized. Occasional, incidental use of these tests must not be expected to show their real potentialities.
3. *The investigation has shown that when the self-instructional tests were used systematically in college courses as an integral part of the teaching method, gains were substantial and sufficiently generalized to improve understanding of a topic as a whole—even help on related topics.* Thus used, the punchboard test brought better work in regular classes,

as shown by higher scores on midterm tests and final examinations (even though used in only about a third of the course) than made by comparable sections of the same course not employing the punchboard. The device was found especially valuable with superior students. Sections of such students, using a series of self-instructional tests systematically covering a course, did superior work with saving of 60 percent or more in number of class hours. There was thus marked reduction in amount of instruction needed, and the students were able to save sufficient time so that they could take extra work, many finishing their education sooner in consequence.

   It appeared that substantial use of carefully planned comprehensive series of instructional tests could have advantages, over the same time in discussion, somewhat comparable to the advantages of an objective test over an essay-type examination or an oral quiz. More questions could be asked of every student in a given time, thus covering a topic more systematically and adequately. The right answers in the tests could be very carefully phrased to be adequate. The wrong answers could bring up and then dispose of common misconceptions. Such helps as page references after each question could guide review. The self-scoring device brings all these values immediately and incisively.

4. This paper reports results with the punchboard. It was found simple and convenient even for student self-use and could readily be made yet more compact and convenient. However, *the total project has shown that there are various promising means for automatic scoring and self-instruction.* A "chemocard" was tried. It was 3" x 5" and looked much like the face of the punchboard. Across the top was space for the student's name, the date, and like information. Then followed two columns of 15 rows of four circles, the rows being numbered from 1 to 30. It could thus be used with 30 questions, four-choice tests. Each student was given a cheap fountain pen filled with a special red ink with which he wrote his name and other information at the top. He then marked through the first circle of the first row if he considered the first alternative for the first question to be right. If it was, the mark immediately turned gray (as a

result of an invisible chemical printed in the "right" circle); if the mark did not turn, the student tried another circle in the first row, continuing until the right answer was found. Total errors would be the total of unchanged red marks. The chemocard obviously operates to test and teach in the same way as the punchboard, but the materials are simpler.[3]

A key machine developed on the basis of the writer's 1926 device not only indicates errors and requires finding of rights before continuing, but also keeps a cumulative count of errors and can readily be used for repeated runs through the test to reach mastery. Yet another mechanism, based on his 1927 apparatus, adds to the above features, in that on successive runs through a test only those questions are returned for further drill on which a mistake was made the previous time through. Such selective review according to each student's need should be yet more effective.[4]

In short, "human engineering" can aid educational and training programs by test-teach devices of various types. The major purpose of this project has been to evidence the value of the basic idea, as illustrated by the punchboard, and to determine ways of using such a device in order to improve substantially instruction, or training.

---

[3] It is believed possible to print right circles with an invisible chemical so that any pen and ink can be used and the right answer will at once blur as in soft blotting paper. The idea of an invisible chemical to indicate correctness of response was developed by the writer's former student, the late Hans Peterson and his brother, John, of Kansas State College at Manhattan.

[4] This device would seem far superior to the Navy's well-known Automatic Rater. About the size of a comptometer, it requires that the right answer be found to each question before going on to the next, returns a card series to starting position by simply pulling a lever, and progressively selects for review the questions causing each student trouble, as mentioned above.

Certain Special Factors Involved in the Law of Effect

AVERY LIVINGSTON STEPHENS

THE INVESTIGATION used two devices which make multiple-choice tests automatically self-scoring and self-instructional. The "punchboard" has already been described (Pressey, 1950).

. . . .

The Drum Tutor is shaped like a small box measuring 3½" x 4½" x 5". Toward the front are four keys, a window showing the number of the test question to be answered, and an error window showing a cumulative count of the errors. For the question, the number of which shows in the item window, a subject who is using the machine with a test pushes the key corresponding to the choice he believes to be correct. If his choice is wrong, the error count increases by one and the item window remains the same; this indicates that he is to try again with another alternative response to that question. If his choice is correct, the item window shows the number of the next question and the error count does not change. Thus each individual knows immediately whether each response is right or wrong and has before him a cumulative count of his errors.

More than 1500 students who were in various psychology classes at The Ohio State University were subjects for this investigation. All practice tests consisted of 30 multiple-choice, four-alternative, questions. Four kinds of material were used: with the punchboards—nonsense syllables, easy Russian vocabulary, and hard English vocabulary; and with the Drum Tutor—these same three plus subject-matter tests in a course in educational psychology.

The first issue concerned the effects of immediate knowledge of correctness of response and guidance to the right answer in relation to the meaningfulness of the material being learned.

The original form of this article was printed in *Abstracts of Doctoral Dissertations*, No. 64, The Ohio State University Press, 1953. The present version, slightly condensed, was edited by S. L. Pressey, under whose direction the study was done.

Tests of nonsense syllables, easy Russian vocabulary, or hard English vocabulary were submitted to classes which used the punchboards, so that each student knew as he punched in an answer whether it was right; and if not, he tried other alternatives until the correct one was found. (The method here is called "retained," following the terminology of other investigators, because the questions are retained before the learner until he finds the right answer.) Each practice test was taken three times, and then a check test different in form and order of items was taken to assure that any learning was independent of the order or other rote factors. . . . The punchboards were also used for collecting data in the test-as-test situation (each subject makes only one response to each item without obtaining any information as to its rightness or wrongness) and in the vanishing situation, used by Thorndike (each subject makes one response to each item, gaining information of its rightness or wrongness but not responding again to that item until the next time through the series of questions). The retained situation was most effective. Thus with the English material in the test-as-test situation there was no progress from Trial 1 to Trial 3; in the vanishing situation the number right increased from eight to 17 and in the retained situation from eight to 24. The end test showed seven right in the first situation, 16 in the second, and 27 right in the retained situation. Also, the least meaningful material showed least progress from Trials 1 to 3: in the retained situation, nonsense rights increased from six to 16, but Russian rights increased from nine to 27.

An advantage of the retained situation is that as many as three errors can be made on a given question (a total of 90 on a 30-item, four-choice test), and therefore a highly sensitive measure of progress may be obtained. The initial error scores were quite similar for nonsense, Russian, and English: 44, 40, and 42. However, on Trial 2, the errors for both the Russian and English are substantially less (12 and 16) than for the nonsense syllables (32); that is, errors were eliminated more rapidly with the more meaningful materials.

Table 1 shows most persistence of correct responses and reduction of errors with the more meaningful materials and, with these, in the retained situation. All errors on Trial 1 were divided into the number of errors made by the same person on the same item on Trial 2. Repetitions in the test-as-test situation served as an empirical basis for comparisons, since no indication of correctness of responses is given. Clearly, an error is much less likely to recur when a subject has been informed of its wrongness. Thorndike's statements (Thorndike, 1932; Hilgard, 1948) to the contrary are evidently wrong, when empirical repetition data are used as a basis for comparison. Also, correction of errors is greater with more meaningful materials.

Table 1. COMPARISON OF MATERIALS OF VARYING MEANINGFULNESS IN THE (A) PERSISTENCE OF CORRECT RESPONSES, (B) PERSISTENCE OF SAME ERRORS, AND (C) CORRECTION OF ERRORS IN THE TEST-AS-TEST, VANISHING, AND RETAINED SITUATIONS

		NONSENSE	RUSSIAN	ENGLISH
Percent of responses right first response on Trial 1 and also right first response on Trial 2	Test-as-test	65%	55%	72%
	Vanishing	66	85	71
	Retained	50	85	81
Percent of repetition of same error from Trial 1 to Trial 2	Test-as-test	54	50	65
	Vanishing	34	20	20
	Retained	37	22	17
Percent of times one error was made on Trial 1, and this was corrected on first try on Trial 2	Test-as-test	18	18	12
	Vanishing	21	30	32
	Retained	43	74	79

. . . . . .

Special test arrangements might conceivably improve learning. To investigate this possibility, an inside-alternates test (each of the wrong alternatives for each item appears as a right choice for another item) was compared with the usual type in which none of the alternatives, right or wrong, appears more than once. No difference was found on an end test, after three times through a practice test, for either nonsense syllables or Russian. For the Russian, a special grouping of the inside alternatives (each right alternative appeared as a wrong alternative for the three items immediately following), designed to assist individuals in finding the right choices, greatly reduced the number of practice-test errors but not of errors on a recall test. A Russian paired-inside-alternates test (between two Russian words, one at the beginning and the other at the end of each line, were the two English words meaning the same as the Russian terms; the student was asked to indicate the meaning of the last Russian word) was compared with a two-choice test on the same Russian words. The last showed 19 errors on a recall-end test after three practice trials, while the special test showed only 13 errors (the difference is significant at the 1 percent level of confidence).

Credence is given to the hypothesis that special test arrangements which greatly reduce the number of wrong responses and increase potential right associations on practice trials are beneficial for learning. Thus the special Russian test had the meanings of both Russian words in each line—and each occurred twice in the test.

The Drum Tutor would seem to have two advantages over the punchboard: (a) Each student constantly has before him a cumulative count of his errors and a record of the number of errors for each trial, and (b) repeated runs through a test are easy because the students quickly learned to reset their machines for the next trial. Comparison of Drum Tutor with punchboard for three practice trials showed the first advantage of no importance. Repeated runs, however, were found valuable. End test scores after one week were increasingly better in proportion to increasing numbers of practice trials. Moreover, no harm was done by repeating the same test: Three times through the same practice test gave as good end test results as using three tests which changed the order of the items and pattern of the right alternatives; nor did distributed practice yield better end test results than the massed practice facilitated by the Drum Tutor. Thus the convenience of repeated testing of the Drum Tutor seemed a real advantage.

The Drum Tutor was therefore so used in a section of an educational psychology course with 11 practice tests on the course material. The students continued trials with each test to one perfect or near-perfect trial. This group was found to achieve better scores on the three midterms and the final examinations than control groups not using these tests or devices, not only with items repeated from the practice tests, but also with new items and with differently phrased items about the same topics. The Drum Tutor section (even though inferior to all other sections in general ability and previous college record) did, in general, as well on the examinations as several groups using the punchboard for only one trial with these same practice tests and did better on items repeated from the practice tests. The Drum Tutor was unanimously favored by those using it; some wished it had been used more frequently.

It appears to have demonstrated that such devices for immediate self-scoring may greatly expedite research regarding the learning process. Data such as that which Thorndike and his associates gathered laboriously by detailed individual procedure were quickly and conveniently obtained from entire classes. It would appear that these devices were valuable research instruments. Also, as a project in human engineering, the findings suggest ways of increasing efficiency in learning or training by special organizations of practice tests or by facilitated retesting.

PART III

Skinner's Teaching Machines And Programming Concepts

IN HIS early studies of the learning process, E. L. Thorndike had a dual concern with two related problems. The first of these was the analysis, through experiment and theorization, of fundamental factors that govern the modification of behavior through learning. The second was a practical concern with methods for improving the effectiveness of educational procedures and materials. This dual interest has been reflected in the frequent citations of Thorndike's work in educational textbooks and journals as well as in an expanding experimental literature concerned with laboratory studies of variables that affect learning and with theories describing the operation of these variables. The latter work has been based on observations of the modification of behavior, both in human subjects and in infrahuman species, in simplified learning situations in the experimental laboratory.

However, since Thorndike's time most of this experimental work and accompanying development of learning theory has proceeded with little explicit reference to the practical problems of education. It thus became something of a rarity to find a single individual actively concerned, as was Thorndike, both with development of the basic science of learning and with its technological applications to educational practice. There were, of course, exceptions. For example, as a theorist, E. R. Guthrie discussed his notions of basic mechanisms in learning primarily by illustrating their operation in a variety of practical learning situations. On the other hand, as an educational psychologist, S. L. Pressey sought, as we have seen, to implement principles of learning in classroom teaching devices. Also, beginning with World War II, an increasing number of experimental psychologists began to apply their concepts and techniques to the problems of military training. But, by and large, most educators and experimental psychologists went their separate ways with only sporadic concern for each other's problems and the possibilities for fruitful interaction. Experimentation in classroom procedures and the laboratory studies by experimental psychologists were reported in separate journals, with surprisingly little overlap (see the cogent discussion by Melton, 1959a).

The possibility of a more direct educational application of concepts developed in the experimental laboratory is advocated by B. F. Skinner in his 1954 paper, "The Science of Learning and the Art of Teaching." In this paper, reprinted here as the first of the articles in Part III, Skinner proposes that his extensive work on the experimental analysis of behavior has direct implications for the teaching process and that its application can be effectively implemented by appropriate instrumentation. In this article he points out the relevance to education of experimental laboratory work on the control and modification of behavior and illustrates his ideas by describing an experimental teaching device. This mechanical device is designed to present a carefully sequenced set of materials and to reinforce a student's response at each step of a program by which he can progress toward the desired behavioral capabilities.

Following Skinner's 1954 article, an intensive program was pursued at Harvard under his guidance to develop and put to work the ideas he had expressed. This work is represented in a number of papers which comprise Part III of this volume and in several additional papers abstracted in Appendix I.

The first of the later papers included here is by Douglas Porter. It is a survey of some past work on mechanical teaching aids. In this paper Porter makes an important distinction among devices which provide stimulus functions only (such as usual forms of audio-visual aids), those which provide response apparatus only (such as Pressey's punchboard), and those which provide for both stimulus and response functions. As examples of the last class, Porter cites Pressey's early self-testing and teaching devices and the first machines developed in Skinner's laboratory, pointing out the similarities and differences in the two approaches. Two important innovations made by Skinner are brought out here—the use of constructed answers rather than multiple-choice responding and the carefully organized programming of subject matter. Porter evaluates four previous experimental studies reported in Part II of this volume.

The next paper included here is by L. E. Homme (1957). This gives a brief report of work then in progress at Harvard on programs for teaching various subject matters and of the rationale underlying this work. During 1957 Skinner also prepared several unpublished reports on techniques of programming materials for teaching machines, for the Army-sponsored Human Resources Research Office of George Washington University. These are among the papers from Skinner's laboratory which are abstracted in Appendix I. A brief description of Skinner's devices and concepts of programming was also given in an unpublished but rather widely cited memorandum prepared by W. D. Edwards for the Air Force on the basis of an interview with Skinner early in 1956 (see Appendix I).

A comprehensive report of the work at Harvard on teaching machines and the programming of materials used with them was prepared in 1958 as a report to the Fund for the Advancement of Education. Skinner summarized the first part of this report in a symposium paper at the 1958 meetings of the American Psychological Association and shortly thereafter published it in *Science* (October 24, 1958, issue). This major article, reproduced in full as the third paper in Part III, attracted wide attention to the potentialities of "teaching machines." It also focused attention on the *"programming"* of detailed, carefully ordered learning sequences by which complex behavioral repertoires could be shaped through successive approximations.

The remainder of this 1958 report, not hitherto published for general distribution, comprises the fourth paper in this series. This paper, by Skinner and Holland, discusses some of the specific characteristics of teaching machine design and variables in the programming of verbal knowledge, and it describes procedures and techniques useful in constructing programs. Some related considerations and additional examples of techniques also have been noted by Skinner in a later paper (1959b) abstracted in Appendix I.

The next two papers are concerned with the teaching of foreign languages. In the first of these C. B. Ferster and Stanley M. Sapon report on an initial attempt to utilize Skinner's concepts in providing individual instruction in first-semester college German. An interesting aspect of their report is the use of a simple paper mask instead of a more complex machine for presenting the programmed sequence of instructional steps. In the second paper on language instruction Porter discusses the functions of various kinds of visual and audio devices for use in second-language learning as well as the application of reinforcement and programming concepts to the requirements of language-teaching devices.

In another article Porter presents a general overview of underlying principles and classroom applications of programmed learning. This article provides a concise, nontechnical presentation which many individuals have found useful. A more detailed and analytic presentation is provided in the paper by J. G. Holland, who has been associated with Skinner in much of the teaching machine development at Harvard. This article illustrates several devices for use at different levels of education and discusses seven principles of programming which represent the "application of methods for behavioral control in developing programs for teaching." The important principle of program improvement through empirical tryout and revision is illustrated by a comparison of original and modified versions of a program for a college course in psychology. (Excerpts from this program and the distribution of answers given by students have been previously reported by Holland [1959]; see Appendix I.)

The two final papers in Part III are concerned with programs written for teaching at the high-school and preschool levels. The first of these papers, by Susan R. Meyer, reports an experiment comparing various response-correction procedures in a portion of a program for vocabulary development in secondary schools. An especially interesting aspect of this paper is the author's detailed analysis of the functions of various forms of prompting used in programming. The last of the papers presented here is by W. Hively. This paper describes an experimental apparatus and technique for automatic teaching of prereading skills to nursery-school children. The apparatus is so devised that the child can learn to make discriminations and to match objects, while receiving appropriate reinforcement for correct responses without the intervention of the teacher.

Other contributions from the work at Harvard are abstracted in Appendix I. These include a summary of some of the extensive work done by Porter (1959) on the programming of spelling for elementary-school grades and a brief report by Meyer

(1959) on her work in teaching elementary arithmetic by machine-presented programs; Also included in Appendix I are abstracts of two other general discussions on the implications of Skinner's analysis of behavior for educational practice, by Hively (1959) and Israel (1959).

It is the work of Skinner and his associates that has caught the imagination of educators and psychologists throughout the country and that has provided the main stimulus for the surge of interest that has led to the publication of the present volume. Some examples of the numerous projects that have been begun elsewhere in the country as a consequence of Skinner's leadership are presented in Part V. A penetrating discussion of some more general implications of the needs for interaction between the science of learning and the technology of educational methods has been presented by Melton (1959a). Some of the editor's views on the implications of research and development on teaching machines and programmed learning for the development of an instructional technology based on the science of learning are presented in the concluding section of this volume [p. 563 ff.].

The Science of Learning And the Art of Teaching

B. F. SKINNER

SOME PROMISING advances have recently been made in the field of learning. Special techniques have been designed to arrange what are called "contingencies of reinforcement"—the relations which prevail between behavior on the one hand and the consequences of that behavior on the other—with the result that a much more effective control of behavior has been achieved. It has long been argued that an organism learns mainly by producing changes in its environment, but it is only recently that these changes have been carefully manipulated. In traditional devices for the study of learning—in the serial maze, for example, or in the T-maze, the problem box, or the familiar discrimination apparatus—the effects produced by the organism's behavior are left to many fluctuating circumstances. There is many a slip between the turn-to-the-right and the food-cup at the end of the alley. It is not surprising that techniques of this sort have yielded only very rough data from which the uniformities demanded by an experimental science can be extracted only by averaging many cases. In none of this work has the behavior of the individual organism been predicted in more than a statistical sense. The learning processes which are the presumed object of such research are reached only through a series of inferences. Current preoccupation with deductive systems reflects this state of the science.

Recent improvements in the conditions which control behavior in the field of learning are of two principal sorts. The "law of effect" has been taken seriously; we have made sure that effects *do* occur and that they occur under conditions which are optimal

---

This article is reprinted from the *Harvard Educational Review*, Vol. 24, No. 2, 1954. It is based on a paper presented at a conference on current trends in psychology at the University of Pittsburgh in March 1954, and it also appeared in *Current Trends in Psychology and the Behavioral Sciences*, published by the University of Pittsburgh Press (1955). Figure I (p. 113) was included in Skinner's *Cumulative Record* (New York: Appleton-Century, 1958).

for producing the changes called learning. Once we have arranged the particular type of consequence called a reinforcement, our techniques permit us to shape the behavior of an organism almost at will. It has become a routine exercise to demonstrate this in classes in elementary psychology by conditioning such an organism as a pigeon. Simply by presenting food to a hungry pigeon at the right time, it is possible to shape up three or four well-defined responses in a single demonstration period—such responses as turning around, pacing the floor in the pattern of a figure-8, standing still in a corner of the demonstration apparatus, stretching the neck, or stamping the foot. Extremely complex performances may be reached through successive stages in the shaping process, the contingencies of reinforcement being changed progressively in the direction of the required behavior. The results are often quite dramatic. In such a demonstration one can *see* learning take place. A significant change in behavior is often obvious as the result of a single reinforcement.

A second important advance in technique permits us to maintain behavior in given states of strength for long periods of time. Reinforcements continue to be important, of course, long after an organism has learned *how* to do something, long after it has acquired behavior. They are necessary to maintain the behavior in strength. Of special interest is the effect of various schedules of intermittent reinforcement. Charles B. Ferster and the author are currently preparing an extensive report of a five-year research program, sponsored by the Office of Naval Research, in which most of the important types of schedules have been investigated and in which the effects of schedules in general have been reduced to a few principles. On the theoretical side we now have a fairly good idea of why a given schedule produces its appropriate performance. On the practical side we have learned how to maintain any given level of activity for daily periods limited only by the physical exhaustion of the organism and from day to day without substantial change throughout its life. Many of these effects would be traditionally assigned to the field of motivation, although the principal operation is simply the arrangement of contingencies of reinforcement. (See also Skinner, 1953b).

These new methods of shaping behavior and of maintaining it in strength are a great improvement over the traditional practices of professional animal trainers, and it is not surprising that our laboratory results are already being applied to the production of performing animals for commercial purposes. In a more academic environment they have been used for demonstration purposes which extend far beyond an interest in learning as such. For example, it is not too difficult to arrange the complex contingencies which produce many types of social behavior. Competition is exemplified by two pigeons playing a modified game of ping-pong. The pigeons drive the ball back and forth across a small table by pecking at it. When the ball gets by one pigeon, the other is reinforced. The task of constructing such a "social relation" is probably completely out of reach of the traditional animal trainer. It requires a carefully designed program of gradually changing contingencies and the skillful use of schedules to maintain the behavior in strength. Each pigeon is separately prepared for its part in the total performance, and the "social relation" is then arbitrarily constructed. The sequence of events leading up to this stable state are excellent material for the study of the factors important in nonsynthetic social behavior. It is instructive to consider how a similar series of contingencies could arise in the case of the human organism through the evolution of cultural patterns.

Cooperation can also be set up, perhaps more easily than competition. We have trained two pigeons to coordinate their behavior in a cooperative endeavor with a precision which equals that of the most skillful human dancers. In a more serious vein these techniques have permitted us to explore the complexities of the individual organism and to analyze some of the serial or coordinate behaviors involved in attention, problem solving, various types of self-control, and the subsidiary systems of responses within a single organism called "personalities." Some of these are exemplified in what we call multiple schedules of reinforcement. In general, a given schedule has an effect upon the rate at which a response is emitted. Changes in the rate from moment to moment show a pattern typical of the schedule. The pattern may be as simple as a constant rate of responding at a given value, it may be a gradually accelerating rate between certain extremes, it may be an abrupt change from not responding at all to a given stable high rate, and so on. It has been shown that the performance characteristic of a given schedule can be brought under the control of a particular stimulus and that different performances can be brought under the control of different stimuli in the same organism. At a recent meeting of the American Psychological Association, Dr. Ferster and the author demonstrated a pigeon whose behavior showed the pattern typical of "fixed-interval" reinforcement in the presence of one stimulus and, alternately, the pattern typical of the very different schedule called "fixed ratio" in the presence of a second stimulus. In the laboratory we have been able to obtain performances appropriate to *nine* different schedules in the presence of appropriate stimuli in random alternation. When Stimulus 1 is present, the pigeon executes the performance appropriate to Schedule 1. When Stimulus 2 is present, the pigeon executes the performance appropriate to Schedule 2. And so on. This result is important because it makes the extrapolation of our laboratory results to daily life much more plausible. We are all constantly shifting from schedule to schedule as our immediate environment changes, but the dynamics of the control exercised by reinforcement remain essentially unchanged.

It is also possible to construct very complex *sequences* of schedules. It is not easy to describe these in a few words, but two or three examples may be mentioned. In one experiment the pigeon generates a performance appropriate to Schedule A where the reinforcement is simply the production of the stimulus characteristic of Schedule B, to which the pigeon then responds appropriately. Under a third stimulus the bird yields a performance appropriate to Schedule C where the reinforcement in this case is simply the production of the stimulus characteristic of Schedule D, to which the bird then responds appropriately. In a special case, first investigated by L. B. Wyckoff, Jr., the organism responds to one stimulus where the reinforcement consists of the *clarification* of the stimulus controlling another response. The first response becomes, so to speak, an objective form of "paying attention" to the second stimulus. In one important version of this experiment, as yet unpublished, we could say that the pigeon is telling us whether it is "paying attention" to the *shape* of a spot of light or to its *color*.

One of the most dramatic applications of these techniques has recently been made in the Harvard Psychological Laboratories by Floyd Ratliff and Donald S. Blough, who have skillfully used multiple and serial schedules of reinforcement to study complex perceptual processes in the infrahuman organism. They have achieved a sort of psycho-physics without verbal instruction. In a recent experiment by Blough, for example, a pigeon draws a detailed dark-adaptation curve showing the characteristic breaks of rod and cone vision. The curve is recorded continuously in a single experimental period and is quite comparable with the curves of human subjects. The pigeon behaves in a way which, in the human case, we would not hesitate to describe by saying that it adjusts a very faint patch of light until it can just be seen.

In all this work, the species of the organism has made surprisingly little difference. It is true that the organisms studied have all been vertebrates, but they still cover a wide range. Comparable results have been obtained with pigeons, rats, dogs, monkeys, human children, and most recently—by the author in collaboration with Ogden R. Lindsley—with human psychotic subjects. In spite of great phylogenetic differences, all these organisms show amazingly similar properties of the learning process. It should be emphasized that this has been achieved by analyzing the effects of reinforcement and by designing techniques which manipulate reinforcement with considerable precision. Only in this way can the behavior of the individual organism be brought under such precise control. It is also important to note that through a gradual advance to complex interrelations among responses, the same degree of rigor is being extended to behavior which would usually be assigned to such fields as perception, thinking, and personality dynamics.

From this exciting prospect of an advancing science of learning, it is a great shock to turn to that branch of technology which is most directly concerned with the learning process—education. Let us consider, for example, the teaching of arithmetic in the lower grades. The school is concerned with imparting to the child a large number of responses of a special sort. The responses are all verbal. They consist of speaking and writing certain words, figures, and signs which, to put it roughly, refer to numbers and to arithmetic operations. The first task is to shape up these responses—to get the child to pronounce and to write responses correctly, but the principal task is to bring this behavior under many sorts of stimulus control. This is what happens when the child learns to count, to recite tables, to count while ticking off the items in an assemblage of objects, to respond to spoken or written numbers by saying "odd," "even," "prime," and so on. Over and above this elaborate repertoire of numerical behavior, most of which is often dismissed as the product of rote learning, the teaching of arithmetic looks forward to those complex serial arrangements of responses involved in original mathematical thinking. The child must acquire responses of transposing, clearing fractions, and so on, which modify the order or pattern of the original material so that the response called a solution is eventually made possible.

Now, how is this extremely complicated verbal repertoire set up? In the first place, what reinforcements are used? Fifty years ago the answer would have been clear. At that time educational control was still frankly aversive. The child read numbers, copied numbers, memorized tables, and performed operations upon numbers to escape the threat of the birch rod or cane. Some positive reinforcements were perhaps eventually derived from the increased efficiency of the child in the field of arithmetic, and in rare cases some automatic reinforcement may have resulted from the sheer manipulation of the medium —from the solution of problems or the discovery of the intricacies of the number system. But for the immediate purposes of education the child acted to avoid or escape punishment. It was part of the reform movement known as progressive education to make the positive consequences more immediately effective, but anyone who visits the lower grades of the average school today will observe that a change has been made, not from aversive to positive control, but from one form of aversive stimulation to another. The child at his desk, filling in his workbook, is behaving primarily to escape from the threat of a series of minor aversive events—the teacher's displeasure, the criticism or ridicule of his classmates, an ignominious showing in a competition, low marks, a trip to the office "to be talked to" by the principal, or a word to the parent who may still resort to the birch rod. In this welter of aversive consequences, getting the right answer is in itself an insignificant event, any effect of which is lost amid the anxieties, the boredom, and the aggressions which are the inevitable by-products of aversive control (Skinner, 1953a).

Secondly, we have to ask how the contingencies of reinforcement are arranged. When is a numerical operation reinforced as "right"? Eventually, of course, the pupil may be able to check his own answers and achieve some sort of automatic reinforcement, but in the early stages the reinforcement of being right is usually accorded by the teacher. The contingencies she provides are far from optimal. It can easily be demonstrated that, unless explicit mediating behavior has been set up, the lapse of only a few seconds between response and reinforcement destroys most of the effect. In a typical classroom, nevertheless, long periods of time customarily elapse. The teacher may walk up and down the aisle, for example, while the class is working on a sheet of problems, pausing here and there to say right or wrong. Many seconds or minutes intervene between the child's response and the teacher's reinforcement. In many cases—for example, when papers are taken home to be corrected—as much as 24 hours may intervene. It is surprising that this system has any effect whatsoever.

A third notable shortcoming is the lack of a skillful program which moves forward through a series of progressive approximations to the final complex behavior desired. A long series of contingencies is necessary to bring the organism into the possession of mathematical behavior most efficiently. But the teacher is seldom able to reinforce at each step in such a series because she cannot deal with the pupil's responses one at a time. It is usually necessary to reinforce the behavior in blocks of responses—as in correcting a work sheet or page from a workbook. The responses within such a block must not be interrelated. The answer to one problem must not depend upon the answer to another. The number of stages through which one may progressively approach a complex pattern of behavior is therefore small and the task so much the more difficult. Even the most modern workbook in beginning arithmetic is far from exemplifying an efficient program for shaping up mathematical behavior.

Perhaps the most serious criticism of the current classroom is the relative infrequency of reinforcement. Since the pupil is usually dependent upon the teacher for being right and since many pupils are usually dependent upon the same teacher, the total number of contingencies which may be arranged during, say, the first four years is of the order of only a few thousand. But a very rough estimate suggests that efficient mathematical behavior at this level requires something of the order of 25,000 contingencies. We may suppose that even in the brighter student a given contingency must be arranged several times to place the behavior well in hand. The responses to be set up are not simply the various items in tables of addition, subtraction, multiplication, and division; we have also to consider the alternative forms in which each item may be stated. To the learning of such material we should add hundreds of responses concerned with factoring, identifying primes, memorizing series, using short-cut techniques of calculation, constructing and using geometric representations or number forms, and so on. Over and above all this, the whole mathematical repertoire must be brought under the control of concrete problems of considerable variety. Perhaps 50,000 contingencies is a more conservative estimate. In this frame of reference the daily assignment in arithmetic seems pitifully meager.

The result of all this is, of course, well known. Even our best schools are under criticism for their inefficiency in the teaching of drill subjects such as arithmetic. The condition in the average school is a matter of widespread national concern. Modern children simply do not learn arithmetic quickly or well. Nor is the result simply incompetence. The very subjects in which modern techniques are weakest are those in which failure is most conspicuous, and in the wake of an ever-growing incompetence come the anxieties, uncertainties, and aggressions which, in their turn, present other problems to the school. Most pupils soon claim the asylum of not being "ready" for arithmetic at a given level or, eventually, of not having a mathematical mind. Such explanations are readily seized upon by defensive teachers and parents. Few pupils ever reach the stage at which automatic reinforcements follow as the natural consequences of mathematical behavior. On the contrary, the figures and symbols of mathematics have become standard emotional stimuli. The glimpse of a column of figures, not to say an algebraic symbol or an integral sign, is likely to set off—not mathematical behavior—but a reaction of anxiety, guilt, or fear.

The teacher is usually no happier about this than the pupil. Denied the opportunity to control via the birch rod, quite at sea as to the mode of operation of the few techniques at her disposal, she spends as little time as possible on drill subjects and eagerly subscribes to philosophies of education which emphasize material of greater inherent interest. A confession of weakness is her extraordinary concern lest the child be taught something unnecessary. The repertoire to be imparted is carefully reduced to an essential minimum. In the field of spelling, for example, a great deal of time and energy has gone into discovering just those words which the young child is going to use, as if it were a crime to waste one's educational power in teaching an unnecessary word. Eventually, weakness of technique emerges in the disguise of a reformulation of the aims of education. Skills are minimized in favor of vague achievements — educating for democracy, educating the whole child, educating for life, and so on. And there the matter ends, for, unfortunately, these philosophies do not in turn suggest improvements in techniques. They offer little or no help in the design of better classroom practices.

There would be no point in urging these objections if improvement were impossible. But the advances which have recently been made in our control of the learning process suggest a thorough revision of classroom practices, and, fortunately, they tell us how the revision can be brought about. This is not, of course, the first time that the results of an experimental science have been brought to bear upon the practical problems of education. The modern classroom does not, however, offer much evidence that research in the field of learning has been respected or used. This condition is no doubt partly due to the limitations of earlier research. But it has been encouraged by a too hasty conclusion that the laboratory study of learning is inherently limited because it cannot take into account the realities of the classroom. In the light of our increasing knowledge of the learning process we should, instead, insist upon dealing with those realities and forcing a substantial change in them. Education is perhaps the most important branch of scientific technology. It deeply affects the lives of all of us. We can no longer allow the exigencies of a practical situation to suppress the tremendous improvements which are within reach. The practical situation must be changed.

There are certain questions which have to be answered in turning to the study of any new organism. What behavior is to be set up? What reinforcers are at hand? What responses are available in embarking upon a program of progressive approximation which will lead to the final form of the behavior? How can reinforcements be most efficiently scheduled to maintain the behavior in strength? These questions are all relevant in considering the problem of the child in the lower grades.

In the first place, what reinforcements are available? What does the school have in its possession which will reinforce a child? We may look first to the material to be learned, for it is possible that this will provide considerable automatic reinforcement. Children play for hours with mechanical toys, paints, scissors and paper, noise-makers, puzzles—in short, with almost anything which feeds back significant changes in the environment and is reasonably free of aversive properties. The sheer control of nature is itself reinforcing. This effect is not evident in the modern school because it is masked by the emotional responses generated by aversive control. It is true that automatic reinforcement from the manipulation of the environment is probably only a mild reinforcer and may need to be carefully husbanded, but one of the most striking principles to emerge from recent research is that the *net* amount of reinforcement is of little significance. A very slight reinforcement may be tremendously effective in controlling behavior if it is wisely used.

If the natural reinforcement inherent in the subject matter is not enough, other reinforcers must be employed. Even in school the child is occasionally permitted to do "what he wants to do," and access to reinforcements of many sorts may be made contingent upon the more immediate consequences of the behavior to be established. Those who advocate competition as a useful social motive may wish to use the reinforcements which follow from excelling others, although there is the difficulty that in this case the reinforcement of one child is necessarily aversive to another. Next in order we might place the good will and affection of the teacher, and only when that has failed need we turn to the use of aversive stimulation.

In the second place, how are these reinforcements to be made contingent upon the desired behavior? There are two considerations here—the gradual elaboration of extremely complex patterns of behavior and the maintenance of the behavior in strength at each stage. The whole process of becoming competent in any field must be divided into a very large number of very small steps, and reinforcement must be contingent upon the accomplishment of each step. This solution to the problem of creating a complex repertoire of behavior also solves the problem of maintaining the behavior in strength. We could, of course, resort to the techniques of scheduling already developed in the study of other organisms but in the present state of our knowledge of educational practices, scheduling appears to be most effectively arranged through the design of the material to be learned. By making each successive step as small as possible, the frequency of reinforcement can be raised to a maximum, while the possibly aversive consequences of being wrong are reduced to a minimum. Other ways of designing material would yield other programs of reinforcement. Any supplementary reinforcement would probably have to be scheduled in the more traditional way.

These requirements are not excessive, but they are probably incompatible with the current realities of the classroom. In the experimental study of learning it has been found that the contingencies of reinforcement which are most efficient in controlling the organism cannot be arranged through the personal mediation of the experimenter. An organism is affected by subtle details of contingencies which are beyond the capacity of the human organism to arrange. Mechanical and electrical devices must be used. Mechanical help is also demanded by the sheer number of contingencies which may be used efficiently in a single experimental session. We have recorded many millions of responses from a single organism during thousands of experimental hours. Personal arrangement of the contingencies and personal observation of the results are quite unthinkable. Now, the human organism is, if anything, more sensitive to precise contingencies than the other organisms we have studied. We have every reason to expect, therefore, that the most effective control of human learning will require instrumental aid. The simple fact is that, as a mere reinforcing mechanism, the teacher is out of date. This would be true even if a single teacher devoted all her time to a single child, but her inadequacy is multiplied many-fold when she must serve as a reinforcing device to many children at once. If the teacher is to take advantage of recent advances in the study of learning, she must have the help of mechanical devices.

The technical problem of providing the necessary instrumental aid is not particularly difficult. There are many ways in which the necessary contingencies may be arranged, either mechanically or electrically. An inexpensive device which solves most of the principal problems has already been constructed. It is still in the experimental stage, but a description will suggest the kind of instrument which seems to be required. The device consists of a small box about the size of a small record player. On the top surface is a window through which a question or problem printed on a paper tape may be seen. The child answers the question by moving one or more sliders upon which the digits 0 through 9 are printed. The answer appears in square holes punched in the paper upon which the question is printed. When the answer has been set, the child turns a knob. The operation is as simple as adjusting a television set. If the answer is right, the knob turns freely and can be made to ring a bell or provide some other conditioned reinforcement. If the answer is wrong, the knob will not turn. A counter may be added to tally wrong answers. The knob must then be reversed slightly and a second attempt at a right answer made. (Unlike the flashcard, the device reports a wrong answer without giving the right answer.) When the answer is right, a further turn of the knob engages a clutch which moves the next problem into place in the window. This movement cannot be completed, however, until the sliders have been returned to zero.

The important features of the device are these: Reinforcement for the right answer is immediate. The mere manipulation of the device will probably be reinforcing enough to keep the average pupil at work for a suitable period each day, provided traces of earlier aversive control can be wiped out. A teacher may supervise an entire class at work on such devices at the same time, yet each child may progress at his own rate, completing as many problems as possible within the class period. If forced to be away from school, he may return to pick up where he left off. The gifted child will advance rapidly but can be kept from getting too far ahead either by being excused from arithmetic for a time or by being given special sets of problems which take him into some of the interesting bypaths of mathematics.

The device makes it possible to present carefully designed material in which one problem can depend upon the answer to the preceding and where, therefore, the most efficient progress to an eventually complex repertoire can be made. Provision has been made for recording the commonest mistakes so that the tapes can be modified as experience dictates. Additional steps can be inserted where pupils tend to have trouble, and ultimately the material will reach a point at which the answers of the average child will almost always be right.

If the material itself proves not to be sufficiently reinforcing, other reinforcers in the possession of the teacher or school may be made contingent upon the operation of the device or upon progress through a series of problems. Supplemental reinforcement would not sacrifice the advantages gained from immediate reinforcement and from the possibility of constructing an optimal series of steps which approach the complex repertoire of mathematical behavior most efficiently.

A similar device in which the sliders carry the letters of the alphabet has been designed to teach spelling. In addition to the advantages which can be gained from precise reinforcement and careful programming, the device will teach reading at the same time. It can also be used to establish the large and important repertoire of verbal relationships encountered in logic and science. In short, it can teach verbal thinking. As to content instruction, the device can be operated as a multiple-choice self-rater.

Some objections to the use of such devices in the classroom can easily be foreseen. The cry will be raised that the child is being treated as a mere animal and that an essentially human intellectual achievement is being analyzed in unduly mechanistic terms. Mathematical behavior is usually regarded not as a repertoire of responses involving numbers and numerical operations, but as evidences of mathematical ability or the exercise of the power of reason. It is true that the techniques which are emerging from the experimental study of learning are not designed to "develop the mind" or to further some vague "understanding" of mathematical relationships. They are designed, on the contrary, to establish the very behaviors which are taken to be the evidences of such mental states or processes. This is only a special case of the general change which is under way in the interpretation of human affairs. An advancing science continues to offer more and more convincing alternatives to traditional formulations. The behavior in terms of which human thinking must eventually be defined is worth treating in its own right as the substantial goal of education.

Of course, the teacher has a more important function than to say "right" or "wrong." The changes proposed would free her for the effective exercise of that function. Marking a set of papers in arithmetic—"Yes, 9 and 6 *are* 15; no, 9 and 7 *are not* 18"—is beneath the dignity of any intelligent individual. There is more important work to be done—in which the teacher's relations to the pupil cannot be duplicated by a mechanical device. Instrumental help would merely improve these relations. One might say that the main trouble with education in the lower grades today is that the child is obviously not competent and *knows it* and that the teacher is unable to do anything about it and *knows that too*. If the advances which have recently been made in our control of behavior can give the child a genuine competence in reading, writing, spelling, and arithmetic, then the teacher may begin to function, not in lieu of a cheap machine, but through intellectual, cultural, and emotional contacts of that distinctive sort which testify to her status as a human being.

Another possible objection is that mechanized instruction will mean technological unemployment. We need not worry about this until there are enough teachers to go around and until the hours and energy demanded of the teacher are comparable to those in other fields of employment. Mechanical devices will eliminate the more tiresome labors of the teacher, but they will not necessarily shorten the time during which she remains in contact with the pupil.

A more practical objection: Can we afford to mechanize our schools? The answer is clearly "Yes." The device I have just described could be produced as cheaply as a small radio or phonograph. There would need to be far fewer devices than pupils, for they could be used in rotation. But even if we suppose that the instrument eventually found to be most effective would cost several hundred dollars and that large numbers of them would be required, our economy should be able to stand the strain. Once we have accepted the possibility and the necessity of mechanical help in the classroom, the economic problem can easily be surmounted. There is no reason why the schoolroom should be any less mechanized than, for example, the kitchen. A country which annually produces millions of refrigerators, dishwashers, automatic washing machines, automatic clothes driers, and automatic garbage disposers can certainly afford the equipment necessary to educate its citizens to high standards of competence in the most effective way.

There is a simple job to be done. The task can be stated in concrete terms. The necessary techniques are known. The equipment needed can easily be provided. Nothing stands in the way but cultural inertia. But what is more characteristic of America than an unwillingness to accept the traditional as inevitable? We are on the threshold of an exciting and revolutionary period, in which the scientific study of man will be put to work in man's best interests. Education must play its part. It must accept the fact that a sweeping revision of educational practices is possible and inevitable. When it has done this, we may look forward with confidence to a school system which is aware of the nature of its tasks, secure in its methods, and generously supported by the informed and effective citizens whom education itself will create.

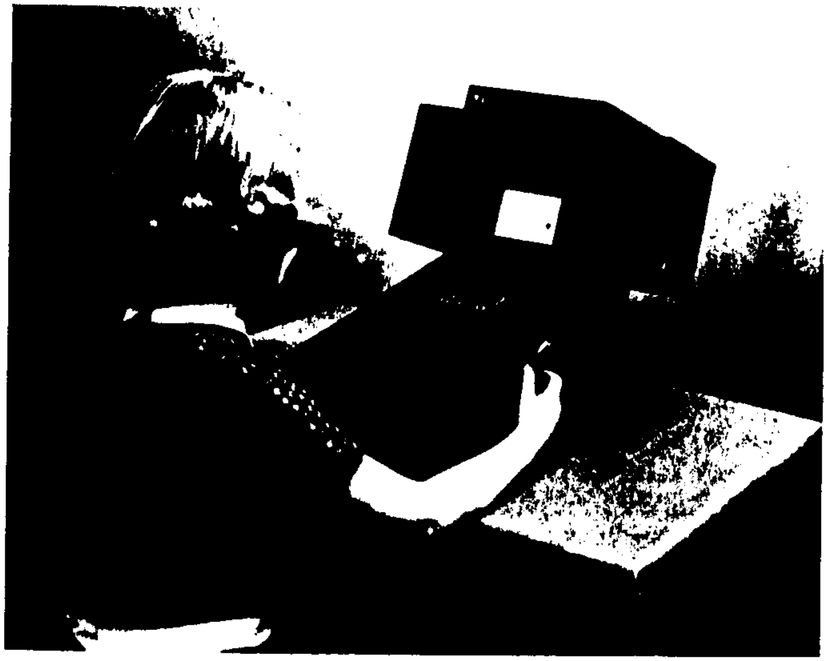


FIG. I. A recent model of a teaching machine for the lower grades. The machine operates on the principles described in the accompanying article. Material is presented in a window with a few letters or figures missing. The pupil moves sliders which cause letters or figures to appear. When an answer has been composed, the pupil turns a crank. If the answer was right, a new frame of material moves into the window and the sliders return to their home position. If the answer was wrong, the sliders return but the frame remains and must be completed again. (This is a later version of the device described in Skinner's 1954 paper.)

A Critical Review of a Portion of The Literature on Teaching Devices

DOUGLAS PORTER

INTRODUCTION

MODERN MASS education has to minister an expanded curriculum to increasing numbers of students with a wide range of individual differences in scholastic ability (Wittich and Schuller, 1953: 4). This fact, coupled with an increasing student-to-teacher ratio (U. S. Department of Health, Education, and Welfare, 1953: 166), has placed schools under pressure of time and numbers. If one of the values of American education is to remain the achievement of individual excellence, some method has to be found to permit more attention to the individual student (Counts, 1952: Ch. 20).

One partial solution which may be proposed is concerned with increasing efficiency of the teaching process itself. With increased efficiency of teaching, the teacher-pupil hours per subject-matter unit could be decreased. As a result, the pressure of insufficient time would be alleviated, allowing more pupils and greater diversity of curriculum to be handled by a smaller number of teachers.

Teaching efficiency might be increased by two closely related measures: (a) the application of some demonstrably valid psychological principles derived from "learning" experiments and (b) the use of teaching aids and devices.

The aim of this paper is to review a portion of the literature concerning the so-called "mechanical teaching machines or devices," which, allegedly, increase the efficiency and effectiveness of teaching operations by making use of certain psychological principles and teacher-laborsaving features.

Most of these teaching devices embody technological advances of the past 30 years. They are, usually, complex electronic, optical, or electro-mechanical devices capable of mediating between the student and what he is studying in some unique manner impossible for the teacher using conventional classroom artifacts alone.

The great variety of mechanical devices reflects their devious origins: Some are the ideas practicing teachers have found to be successful (Berger, 1950; Berger, 1953; Gustafson, 1954; Morley, 1952; Olander, 1955; Torreyson, 1955); others have been sponsored by manufacturers who must envision a large virgin market (Carney, 1951; Schaughency, 1955; Taylor, 1937). Renewed interest in the tachistoscope, an old psychological research instrument, has been brought about by its use in wartime recognition training (Neville, 1943; Renshaw, 1945). Research into the psychology of reading has produced numerous devices (Barbe, 1952; Carter, 1950; Dearborn and Anderson, 1937; Dearborn, Anderson, and Brewster, 1938; Perry and Whitlock, 1954), and certain psychologists have constructed devices which presumably utilize important principles of learning (Dallenbach, 1914; Pressey, 1926; Pressey, 1927; Pressey, 1932; Smith and Tate, 1953).

Although this paper reviews only a small portion of the literature surveyed, a classification of all the devices encountered is felt to be valuable, since: (a) a search of the literature has revealed no other review covering so wide a scope of devices; (b) classification should reveal important differences between devices; (c) the basis of classification will provide a theoretical framework in which literature concerning the devices may be reviewed.

CLASSIFICATION OF MECHANICAL TEACHING AIDS AND DEVICES

In his attempts to facilitate and direct the course of learning, a teacher utilizes various materials as teaching aids. These teaching aids provide the physical stimuli which indicate to a student that certain responses are appropriate. (But for learning to take place, a student has to make the appropriate responses (Guthrie, 1935: 1-11). Here, teaching materials provide the setting in which and with which responses may be perfected.)

Reprinted, with a few editorial changes, from the *Harvard Educational Review*, Vol. 27, No. 2, 1957. An alphabetical listing of types of devices indexed to the references contained in the original article has been deleted in this volume.

Table 1. TYPES OF TEACHING AIDS AND DEVICES

STIMULUS DEVICES	Visual	Models Motion picture Motion picture loop Opaque projector Orthographic projector Printing devices Still film projectors	Stimulus devices, quite simply, do what the subheadings—"Visual," "Auditory," "Audio-Visual," and "Tactile" in Table 1—suggest. They present a student with information about how to make a given response or about when it is appropriate to make the response. They do not provide a setting for the practice of responses under specifiable and controlled circumstances. Thus it may be said that stimulus devices provide learning content without any assurance that a learning process will be carried out.
	Auditory	Phonograph Public address Radio Tape	
	Audio-Visual	Sound motion picture Television	
	Tactile	Braille	
RESPONSE DEVICES	Data manipulators	Abacus Code typewriter Counting devices Desk calculator Typewriter	Response devices, on the other hand, do give a student the opportunity to practice responses. They also may be used to gather, record, and transform coded data about the nature of the responses being made. But response devices provide no stimulus information about the circumstances under which particular behavior is appropriate. Thus responses learned or performed are related to the mechanical properties of the machines and not to any subject-matter content.
	Data gatherers	Alertness indicator Automatic test scorer Classroom communicator Galvanic skin response indicator Recorders	
STIMULUS-RESPONSE DEVICES	Simulators	Add-a-part films and tapes Dearborn reading film (42) Electronic flight simulators	Stimulus devices and response devices considered separately are truncated representations of the teaching-learning situation. However, when both aspects of learning, the presentation of cues and the opportunity for responding, are combined in one device, stimuli and responses are both placed under control of the teaching device and conditions for learning can be made optimum. Stimulus-response devices (or stimulus devices and response devices used in combination) are designed to present a sequence of stimuli (content) and provide the setting in which appropriate responses may be made and rehearsed (process). Thus stimulus-response devices more nearly reproduce the characteristics of teaching which are required for efficient learning than do stimulus devices or response devices.
	Immediate reinforcers	Pressey multiple choice Chemical paper Electric answer board Punchboard Skinner device Trigger squeeze indicator	
	Pacers	Memory drum Metronome Metronoscope Reading accelerators Reading films Tachistoscope	

The use of teaching aids as stimuli distinguished from their use as response materials provides the major basis for a classification of teaching devices. Reference to Table 1 will reveal how various devices have been categorized and some important differences in what may be expected from each class of device.

It should be noted with caution, however, that this classification of devices is not perfect. Many devices fall into one category or another depending upon how they are used. Thus a stimulus device, the motion picture, would have to be classified as a stimulus-response device if provisions were made in the film for audience participation. If a punchboard, classified as a stimulus-response device, were used without any stimulus materials, a student would be able to perfect the punching response, but nothing else. Content, in the sense of a curriculum, would be missing. The punchboard would have to be classified as a response device upon the basis of its use.

On the basis of the distinction between stimulus devices, response devices, and stimulus-response devices, it is suggested that the stimulus type and the response type are not teaching *devices* but teaching *aids*, which must be supplemented by some means, usually a teacher, in order to be effective. The term "teaching device," moreover, should be reserved more correctly for the stimulus-response type of device which is capable of teaching without the mediation of a human teacher. The above usage will be followed in this paper.

SCOPE AND METHOD OF REVIEW

This review is concerned with one of the three stimulus-response types of teaching device: immediate reinforcers. Such restriction of coverage is felt to be warranted because: (a) stimulus-response devices, on theoretical grounds, most adequately meet the requirements of a teaching device suitable for educational use; (b) there already exist numerous reviews of stimulus-type devices (Naval Training Devices Center, 1953; Bortolazzo, 1949; Gibson, 1947; Gibson and others, 1944; Hoban and Van Ormer, 1950; Kinder, 1950; Kinder, 1953; McKown and Roberts, 1949; Wittich and Schuller, 1953); (c) available and educationally relevant literature on response devices is meager; (d) the development of and interest in the immediate reinforcer type of stimulus-response teaching device has been accelerated recently, placing these devices in a controversial position (Edwards, 1956a; Skinner, 1954).

The order in which articles concerning these devices will be evaluated attempts to follow this sequence: (a) a general description of the devices and their use, (b) theoretical issues involved in their use and the experimentation relating to their use, (c) experimental work, (d) endorsement of the devices.

Each article or group of articles concerning a specific type of immediate reinforcer teaching device will be evaluated in terms of six points, based in part upon a paper by Taves and Gross (1952: 111). Application of the six points to individual

articles will depend upon the alleged purpose of the article. In this way, descriptive articles, for instance, will be held to different but equal standards as theoretical papers. The six points of evaluation are:

1. The accurate and complete description of mechanical characteristics of the devices reported
2. The theoretical orientation, which includes an adequate statement of theoretical issues to be investigated and their operational definition
3. The procedures followed, including experimental design and methodology
4. Manipulation of the data gathered, including the legitimate use of statistical techniques
5. A reasonable statement of conclusions which takes into account inherent limitations upon generality of the results
6. The soundness of endorsements recommending use of the devices.

IMMEDIATE REINFORCERS

General Description

Immediate reinforcers are teaching devices which endeavor to alter the course of learning by automatically presenting the student with a "reward" or reinforcement immediately after he has made a correct response. (Technically, a reinforcement is any environmental event the occurrence of which strengthens the behavior it follows. In teaching devices, the reinforcement is a direct consequence of the student's behavior, usually in the form of confirmation of the correct answer.)

These devices all possess three operational characteristics: (a) They present a sequence of problem materials to the student; (b) they provide some means by which the student may record his solution to the problems; and (c) they automatically and immediately indicate correctness of the solution. Each of the devices described below presents some variant of this basic paradigm.

Pressey Multiple-Choice Device

In 1926 S. L. Pressey published the description of "a simple device which gives tests and scores—and teaches." This is the earliest article found concerning immediate reinforcer teaching devices and the first of two articles by Pressey on various models of the same device [A description of the device described by Pressey (1927) was included here in Dr. Porter's article as originally published.]

. . . .

Punchboard Devices

Only six articles concerning this type of device have been found in the literature (Angell and Troyer, 1948; Jensen, 1949; Peterson, 1930; Peterson, 1931; Pressey, 1932; Pressey, 1950). . . . All of the articles present adequate descriptions of the device. The title of the Angell and Troyer paper (1948), however, is misleading. Although they cite the Peterson papers of 18 years earlier . . . which present a similar device, they title their paper "A *New* Self-Scoring Device for Improving Instruction". . . . [A description of the punchboard—see Pressey, 1950—was given here in Dr. Porter's original article.]

. . . .

Chemical Paper

In addition to the work with the punchboards, J. C. [and H. J.] Peterson (1930, 1931) experimented with a chemically treated paper. Pressey mentions this device in his 1950 paper, but no other researchers seem to have been interested. None of the authors give pertinent chemical data for preparation of the materials.

Questions presented with multiple-choice answers were used. The answer space for the correct choice was chemically treated to change color immediately when marked with a special ink. This would inform the student that he had made the correct response. Because the marking ink itself was opaque, a record of all the incorrect responses was also made.

Skinner Device

In a 1954 paper B. F. Skinner presents the description of an immediate reinforcer teaching device. Although Skinner's paper is primarily concerned with development of a theoretical rationale for the application of his teaching device in schools, it would be of value to some readers to have more technical specifications of the device. The thoroughness and detail of description do not match those of Pressey's articles (1926, 1927). [A brief description of Skinner's early device—see Skinner, 1954—was included here by Dr. Porter.]

. . . .

Although different in physical appearance, Skinner's device is essentially a modification of the Pressey device. The Pressey device uses multiple-choice answers, with each of the answers represented by one answer key. This constrains the possible number of answers which may be given by a student and also makes it necessary to present the correct answer. Skinner's modification overcomes the answer constraint by: (a) using as content arithmetic operations producing answers up to four digits, and (b) using four answer keys, each capable of representing the digits 0 through 9. The choice of subject-matter content places inherent constraint upon the number of answers possible, and the positions of the answer keys cover the entire range of content. The consequence of such an arrangement is that students have to compose their answers rather than to choose them from among multiple-choice answers which may be very similar to the correct answer. The advantages of such an arrangement are that it prevents the accidental conditioning of incorrect answers and provides a closer approach to the conditions upon which the answers will be useful in the future. Of course, the design of machines capable of handling materials less constrained than arithmetic becomes very complex but is still feasible.

The other major contribution by Skinner does not reside in the device but in the organization of subject matter presented by the device. Other devices have been used to present groups of questions, relatively unrelated to one another, about some subject-matter area. Skinner presents an organized sequence of materials in which successive problems are dependent upon earlier materials (Skinner, 1954). Successive steps of the subject matter are small, so that if the previous problem has been solved correctly, there is very little chance that the present one will not be also (Skinner, 1954). This raises the frequency of reinforcement to a maximum and reduces the possibility of the aversive consequences of being wrong.

Of course, this type of subject-matter scheduling would be extremely difficult in more amorphous areas than mathematics, but in constructing subject-matter programs, Skinner deals with the problems of curriculum choice, step, and sequence, which are important in all teaching. The criterion for using a particular problem as subject matter is simply that the students must be able to get it right. If the problem cannot be solved by most students, then it is either the wrong problem, at the wrong place in the sequence, or else it is too large a step from the previous problem.

Miscellaneous Devices

H. B. English (1942) describes a device he designed and used during the year 1918 to assist in training soldiers to squeeze rather than jerk a rifle trigger. A tambour was attached to the trigger mechanism of one rifle and connected by rubber tubing to a manometer. Movement of the trigger created a pressure differential which caused liquid in the manometer to move. If the trigger was squeezed properly, the liquid slowly changed level; if the trigger was jerked, the liquid jumped. The author claims great success for this device in teaching poor riflemen to fire correctly. A similar device might be used in the teaching of handwriting to indicate evenness of pencil pressure or smoothness of pencil stroke.

The electric answer board presents a series of questions and answers which the student attempts to match. To check the correctness of a guess, the student touches a contact button next to the answer of his choice with an electric probe. If the answer is correct, a light flashes or a signal sounds. For incorrect choices, nothing happens. Although this is a common children's toy, no articles were found concerning its use in teaching.

Theoretical Issues: Reinforcement

Although there are differences in theoretical explanations of why reinforcement works (Guthrie: Ch. 3, 1935; Hull: Ch. 6, 1943), there can be no doubt that it is effective in the training of laboratory animals. It is also well established that reinforcement is most effective when administered as quickly as possible after a response has been made (Hull: 135-42; 1943). The designers of immediate reinforcer teaching devices have made certain assumptions, however, which are not as well documented. These assumptions, which will be discussed below, are related to the following areas:

1. Generalization of results from animal studies to Homo sapiens
2. Control of the conditions which make knowledge of results reinforcing
3. Type of reinforcement used
4. Quantity of reinforcement
5. Schedule of reinforcement.

A large part of learning-theory research has used the laboratory rat, but there are numerous studies demonstrating the efficacy of reinforcement with human subjects (Woodworth and Schlosberg: 685-93, 1954; Munn: 374-485, 1954). Skinner claims that: "In all this work the species of the organism has made surprisingly little difference. . . . Comparable results have been obtained with pigeons, rats, dogs, monkeys, human children, and most recently . . . human psychotic subjects." (Skinner, 1954). Warren and Brown, 1943, who had children learn a bar-pressing response in a nursery play situation conclude, "The conditioning, extinction, disinhibition and periodic reconditioning of an operant response in children are essentially similar to these phenomena in rats." (Warren and Brown: 203; 1943).

Despite evidence and claims for reinforcement as a fundamental datum of all learning, generalizations from one species to another and from the experimental laboratory to the classroom should be made with caution. The laboratory almost invariably has used solitary subjects, while the classroom provides a chance for complex social interaction. Such interaction may alter the conditions under which certain reinforcements are effective. This is illustrated by recent studies in which Auble and Mech attempted to show the effects of verbal reinforcement in classroom situations (1953 a; 1953 b). There were three groups of subjects: one group was reinforced by verbal praise in front of the entire class; the second group heard praise administered to the first group; and the third group, in another classroom, neither heard nor received verbal praise. When arithmetic achievement of the three groups was compared, the nonpraise group unexpectedly did better than either of the praise-reinforced groups (Auble and Mech, 1953 b). These results may be due to lack of adequately equated groups, accidental praise received by the third group, or by the use of unusual public acclaim which interfered with the reinforcement of arithmetic achievement by disrupting the praised groups' "social structure."

The traditional reinforcers in psychological experimentation have been food, water, and electric shock. There is good reason for their use: all animals will avoid electric shock, hungry or thirsty animals will work for food or water, and it is relatively simple to control the degree of motivation by varying hours of deprivation or intensity of shock. Since it is unethical to shock, starve, or dehydrate children (Berg, 1954), other reinforcers have to be used with teaching devices. Skinner has suggested that exploratory and manipulative activities are reinforcing enough to keep a student at work with a device (Skinner, 1954). Recent experimental studies suggest that this is so (Woodworth and Schlosberg: 685; 1954), but it has also been found that an organism's "curiosity" becomes satiated by prolonged exposure to novel situations (Woodworth and Schlosberg: 671; 1954). Under such circumstances, manipulation and exploration no longer provide adequate reinforcement. To make exploratory activities reinforcing again, the organism has to be removed for a period of time from the formerly novel situation. This restores the effectiveness of reinforcement.

Any teaching device used almost continuously in the classroom would face the same problem of satiation. Unfortunately, none of the authors reviewed deal with this problem. One method of avoiding satiation would be to mete out reinforcement in small doses. Another method would be to have alternative reinforcers available.

Although no author treats in detail the problem of what alternative reinforcers might be used, Skinner suggests several: social approval, desired activities, and as a last resort, aversive stimulation (Skinner, 1954). By including a candy delivery attachment on his device, Pressey (1926) indicates belief that the presenting of sweets can be reinforcing. All of the above events, and many more, can be made contingent upon a student's performance with the teaching device. Whether these events will serve as satisfactory reinforcers will not be known until they are tried (Bijou, 1955).

Most of the authors of articles concerning immediate reinforcer teaching devices are content, however, to side-step the problem of what reinforcers to use. They claim efficacy of their devices is due to "knowledge of results" (Angell and Troyer, 1948; Little: 45, 1934; Pressey, 1950). Indeed, it has been amply demonstrated that knowledge of the correctness of one's own responses improves performance at a task (Woodworth and Schlosberg: 686, 1954). Almost all tasks require some sort of sensory feedback from the environment; this is what knowledge of results provides. However, "knowledge of results" is a descriptive term which refers solely to a particular kind of sensory feedback, not to the reinforcing value of the feedback. The conditions under which this kind of an event attains and maintains reinforcing value are little understood, but if mechanical teaching devices are to be used successfully, educators will have to discover and use the contingencies which are available and most effective in making knowledge of results reinforcing.

Besides type and quantity of reinforcement, the schedule of reinforcement has important effects upon the course of learning. The rate at which a subject will respond, the course of extinction of a response, and the regularity of response all may be brought under the control of reinforcement scheduling (Skinner: 71-77; 1953b). Since distribution of practice depends upon the rate of responding, which is controlled by reinforcement scheduling, rote-learning phenomena (retroactive inhibition, reminiscence, recall) may also be important (McGeoch: Ch. 4; 1942). There is no mention of this in any of the papers reviewed, much less is there any relevant experimental work concerning this aspect of teaching devices.

Theoretical Issues: Transfer

Successful transfer of training from the teaching device to "real" problems is the ultimate criterion of a teaching device's effectiveness. Yet Pressey is the only author to even mention transfer as an area for research. He says:

> The . . . question is . . . whether reactions of the key-pressing type are so different from the reactions to be made in use

> of the materials being learned as to present a serious problem of transfer. . . . Briefly, it may be said that in much school learning there is some analogous problem of transfer; the objection is not believed serious. (Pressey, 1927)

The fact that much school learning is faced with problems of transfer, however, does not minimize the problem. Successful design and use of mechanical devices require the establishment of their validity through the measurement of transfer value.

Illustrative of the type of work needed is a 1939 study by Votaw and Danforth which investigates the effect of method of response upon the validity of multiple-choice tests (Votaw and Danforth, 1939). Equated groups were given the same test but marked their answers differently—by use of a keyed answer sheet, by underlining, and by checking answer spaces. There was found to be a significant difference in scores which depended upon the answer-marking method and was positively correlated with the scores of the subjects on a test of mechanical-transfer ability. The authors conclude, "Apparently the resulting scores represent in part an ability in mechanical-transfer as well as an ability in the subject matter." (Votaw and Danforth: 627; 1939).

It should be expected that what a student learns through use of a mechanical teaching device will depend in part upon characteristics of the mechanism itself and in part upon the subject-matter content which is deliberately presented. Studies similar to the above are needed in order to discover the conditions which will minimize "learning about the machine" and thus maximize learning and transfer of the desired subject matter.

Theoretical Issues: Curriculum

In all of the articles reviewed, immediate reinforcers have been looked upon as teaching and testing devices. No author has mentioned the obvious usefulness of such machines in curriculum research.

Mechanical devices can go far toward eliminating the "teacher effect" in educational investigations. Curriculum materials can be presented to equated groups of students in exactly equivalent manners in order to test the influence of different materials or arrangements of material. The machine removes one heretofore always present source of variation: the use of different teachers, or the same teacher attempting to teach different groups in the same manner.

Experimental Work

There have been found but four experimental articles published on immediate-reinforcer teaching devices. Three of the experiments reported are concerned with punchboard devices (Angell and Troyer, 1948; Little, 1934; Pressey, 1950), one with chemical paper (Peterson, 1930; Peterson, 1931) and one with the Pressey answer-key device (Little, 1934). This certainly does not represent enough experimental work to provide an adequate evaluation of these devices, especially considering the general nature of the work which has been done. All of the experiments attempt to compare the results of mechanical device teaching with "usual" classroom teaching.

The procedure which has been followed is to obtain approximately equated groups of students and expose one group to the usual classroom methods of teaching, whatever they may be, and the other group to mechanical device teaching utilizing the same subject matter. Effectiveness of the two teaching methods is then evaluated by comparing the scores the two groups of students obtain on identical tests.

Such experimentation may indeed show an advantage for one or the other method of teaching, but there is no guarantee that the results obtained can be replicated, for the outcome of those experiments depends upon unspecified parameters of the "usual" classroom situation. As stated by one group of researchers, "The complexity of the teaching-learning process is such that attempts to establish the relative merit of a 'general method of teaching' are likely to prove inconclusive." (Guetzkow, Kelly, and McKeachie: 205; 1954). To be of value, investigations concerning mechanical teaching devices, or any other method of teaching, have to deal with the variables which lie behind the presumed superiority of the method.

The three studies involving the punchboard report superior results from use of that device. But it is maintained here that the results are not due to any demonstrated advantage of the punchboards but to the experimental procedures followed.

Jensen (1949) chose as subjects groups of superior students who wished to accelerate academically. Instead of following the usual classroom procedures in an educational psychology course, they attended a "laboratory" *ad libitum* where they could study, discuss, and test themselves with punchboards. The punchboard tests covered assigned reading in the course, and each question had a page reference next to it so that if a question was missed a student could look up the topic and correct himself. Final course grades of the experimental subjects were superior to those of students attending regular classes.

It is quite obvious that the results obtained may be due to constitution of the experimental group, use of the "laboratory" setting, use of the punchboards, or all three to varying degrees.

In a series of experiments reported by Pressey (1950), the punchboard is used in essentially the same manner as by Jensen to help a student discover what he does not know and thus guide himself to needed further study. In these experiments the test scores of subjects using the punchboard are compared with those of control subjects using conventional multiple-choice tests, or "carefully planned" classroom discussion. Use of the punchboards produced consistently higher scores.

Pressey implies that the difference in scores is due to the "quick incisiveness" of the punchboard (Pressey, 1950). Actually, immediacy of reinforcement may have nothing to do with the results obtained. It is likely that many of the control subjects using conventional multiple-choice tests never went over their test results and thus received no information about what they needed to study. Not having taken *any* tests, it is probably more likely yet that the discussion groups did not find out in what areas they needed study. Thus the fact that the punchboard group received superior test scores may have been because the control subjects received no information concerning their own lack of knowledge; that is, immediacy of reinforcement of the punchboard groups was not responsible for their *superiority*, but lack of knowledge of results in the control groups was responsible for their *inferiority*.

This interpretation is supported by the experiment of Curtis and Woods (1929) in which students who corrected their own tests subsequently received higher scores than students whose tests were corrected by the teacher. Since there was no chance for immediate reinforcement by knowledge of results to take place soon enough to be effective (Hull: 136-40; 1943), the difference in the scores was probably because the group that corrected its own tests was forced to see the mistakes, while the teacher-corrected group was not.

The same sort of criticism is applicable to the studies of Peterson [1931; cf. also 1930] utilizing chemical paper and Little (1934) using both the punchboard and the Pressey device. Methodologically, however, the Peterson and Little researches are superior to those of Pressey and Jensen in both experimental design and statistical analysis. Pressey and Jensen present only average scores for their experimental and control groups. Thus there is no way of determining whether results are significant or not. Peterson and Little, on the other hand, compute the standard error of the difference scores, which does make this information available. Peterson is the only investigator to use other than a simple experimental-control group design. He replicates his experiment, interchanging the experimental and control groups for the replication. In this way, any original differences in the groups are exposed, and the data may be corrected.

It can be stated that the research so far performed shows that immediate reinforcer teaching devices probably aid students in attaining higher test scores than the various alternative classroom procedures with which the devices have been compared. However, it cannot be stated that use of immediate reinforcer devices will produce higher test scores than conventional classroom procedures in general. Nor may it be claimed that the results so far obtained are due to immediacy of reinforcement provided by the devices. A crucial test of the value of immediate reinforcement has to meet the following, so far unattained, criteria: (a) Provide both experimental and control groups with knowledge of results, making sure that both groups receive equivalent information about the correctness of their responses. (b) Reinforce subjects in the experimental group as quickly as possible after a response has been made. (c) Delay reinforcement of the control group.

Literature Endorsements

Authors of articles on immediate reinforcer teaching devices make the following claims for their devices:

1. They are time- and laborsaving for teachers (Little, 1934; Pressey, 1926; Pressey, 1932).
2. They are effective because they provide immediate reinforcement (Pressey, 1950; Skinner, 1954).
3. They are more effective in consistently delivering immediate reinforcement than a teacher can be (Skinner, 1954).
4. They can deal effectively with the range of individual differences present in a classroom (Skinner, 1954).
5. They are able to teach the complex relationships involved in verbal thinking (Skinner, 1954).

There can be no argument about the time- and laborsaving characteristics of devices which present a set of materials to a student and then score and record the results of his endeavors. The teacher is relieved of three tasks: making up and arranging the curriculum, reinforcing the student, and scoring tests.

With individual children operating a machine, the materials presented can be made suitable for the particular child, and the machine can be made to reinforce the child more consistently than a teacher who has to deal with many children. This sort of claim is intuitively obvious. But claims that such devices can be used successfully in all subject-matter areas are premature. Perhaps, hypothetically, teaching devices can be used to teach such complex behaviors as verbal reasoning in social studies, for instance, but at the present stage of technological development they are incapable of dealing with a large enough repertoire of student responses.

This may be the basis for some of the objections voiced against "education by machine" (Pressey, 1926; Skinner, 1954). Critics cannot conceive that complex human endeavors could be analyzed into the precise sort of units suitable for mechanical presentation. Authors of articles on mechanical devices can avoid this sort of resistance if they emphasize the ways in which their devices can help now, at the present stage of development, rather than predicting future successes in less obvious areas.

ACCEPTANCE AND USE OF MECHANICAL TEACHING DEVICES

Although not usually considered as such, almost all classroom paraphernalia are teaching aids and devices (Perry and Whitlock: 6; 1954). Books, blackboards, pencils, and paper have been so long a part of teaching that they are accepted, perhaps, as essential to the teaching-learning process. Newer types of mechanical teaching devices, however, have not been accepted as so essential. As the following quotations show, they tend to be thought of as ancillary to the teaching situation:

> In general, the elaborate mechanical devices should be regarded as a last resort to be used when other methods have failed . . . (Gates, 1947, p. 354). Instruments . . . are not a substitute for a good teacher (Speer, 1951, p. 48). Audio-visual devices can and should, from time to time, be used in the classroom but their proper place is in the language laboratory. Many teachers who believe in the oral approach justly feel that they could get better results if they were given smaller classes and more hours per class rather than more machines (Gaudin, 1952, p. 284).

Authors of the above quotations fail to perceive that conventional classroom paraphernalia and the newer optical, electronic, and electro-mechanical types of device actually have the same basic function: They provide the manipulanda and discriminanda upon which learning is based. Therefore, mechanical devices should not be placed in a special category but should be judged in the same manner as other teaching devices, in terms of their effectiveness. Judging their effectiveness should include determination of the exact conditions under which the devices will do what they are designed to do. This may prevent: (a) overgeneralization of teaching results expected and (b) incorrect and unsuccessful utilization and subsequent casting aside of what may be excellent devices when correctly used.

Table 2. PERCENTAGE OF SCHOOLS USING DEVICES

DEVICE	PERCENTAGE
Motion picture	38.1
Phonograph	38.1
Radio	24.2
Public address	12.3
Opaque projector	11.3
Recorder	9.9
No devices	6.9
Other	2.4

There may be a rational basis for the opinions quoted above, however. Examination of Table 2 will reveal that over 90 percent of the teaching devices in use probably are the stimulus type of teaching aid (Peins: 138; 1951). Such devices provide a student with information and may motivate activity, but stimulus devices do not provide for the rehearsal or practice of behavioral responses appropriate to the "basic" skill areas of schooling, such as reading, spelling, writing, and arithmetic. In not being able to deal with the skill-acquisition areas of schoolwork, stimulus teaching devices are indeed peripheral, and the opinions quoted, valid. It may be that increased use of stimulus-response teaching devices, which do provide for practice of skills, will place mechanical teaching devices in a more prominent classroom role.

The Rationale of Teaching By Skinner's Machines

LLOYD E. HOMME

THE RATIONALE of teaching by machine needs more discussion. This, because it is, in my estimation, the most important technological advance in psychology since the intelligence test and, at the same time, the most likely to be misunderstood.

Most of you are familiar with Skinner's arithmetic teaching machine. Its design requires the student to compose his answer and to receive immediate knowledge of the correctness or incorrectness of his answer. It is intended that the student be led through a series of carefully graded steps, none of which is too difficult to answer, yet each of which is slightly more difficult than the preceding one. This procedure, leading an organism to emit a response which is initially at low strength in his repertoire by reinforcing responses which more and more closely approximate the required one, is known in many laboratories as the reinforcement of successive approximations.

The device which is used for teaching arithmetic in this way is about the size of a portable typewriter. A window, about 3" x 4", exposes a problem, or frame. The student composes his answer by moving sliders on which numbers and operators are mounted. When he is satisfied that he has composed a good answer, he turns a crank. If the answer is correct, the next frame moves into position. If the answer is incorrect, the sliders reset and the student composes another answer. Skinner has devised similar machines for use with adults, the chief difference being that the human, rather than the machine, decides whether a given answer is correct. These machines, together with some sample teaching material, are on display here at this meeting.

This paper was presented at the annual convention of the American Psychological Association in September 1957; previously, Dr. Homme had worked with Dr. Skinner for a year at Harvard. This paper was written during the time Dr. Homme was on the staff of the Department of Psychology of the University of Pittsburgh; at present, he is president of Teaching Machines, Inc., Albuquerque, New Mexico.

Now, I ask you to put out of your minds for a moment the contrivance I have described and consider the material which is presented by it to the student. The reluctance of people to separate these two features of the venture may be the reason that the most frequent characterization we have heard in discussing this project with visitors during the past year is that machine teaching must lead to a kind of mechanical or rote learning. (I needn't add that this is offered in derogation.) I must confess that I am not quite sure what people mean when they say this. The psychologist, I believe, is succumbing to stimulus generalization; he examines the device and reacts to it the way he would to a memory drum. Actually, the two have almost nothing in common. If one considers the way material is presented, there is no similarity.

Both memory drums and teaching machines have the goal of strengthening verbal responses. But there are *many* variables which control the probability of a verbal response. When a high probability of response to stimulation is solely due to the learner's having given the same response before, we call this "rote learning." This is part of what goes on in memory drum learning, and this is what we assiduously avoid in programming machine material. In the material we have prepared so far—Susan Meyer for arithmetic [and] I for elementary college physics—there are *zero* repetitions; exactly the same stimulation never occurs twice. Our goal is to give the subject the largest possible number of reasons for making a response other than the fact that he has made it before. Our rationale is that, to the extent that a student can give reasons for making a response in a problem situation, he has knowledge about it or understands it. This is not easy to bring about; but if we fail, the student, recording his errors as he goes, tells us quite precisely where we fail and where we should try other techniques. This brings me to what I consider the most significant feature of self-teaching by successive approximations. (I use this descriptive phrase rather than "machine-teaching" because, obviously, the verbal program may be presented to the student in a number of ways—not only by machine.)

We have all known teachers in this world whose skill was a subject of wonder to their colleagues and students. But their skill has died with them, so to speak. It is extremely difficult to discriminate, not to mention verbalize, the exact techniques which made them successful. This fact has led intelligent people to characterize teaching as an art and to say, "Good teachers are born, not made."

The point is that learning to teach is an extremely difficult business. Why is it such a difficult skill to learn? Approached from the standpoint of the shaping of behavior, the reasons are clear. For a skill to develop, differential reinforcement contingencies must be precise and frequent. Only if these conditions are met will useless behavior drop out and necessary behavior get strengthened. When we examine the teaching situation, we see that the teacher is in an extremely unfortunate position in this respect. Under the best circumstances differential reinforcement is based on large amounts of verbal behavior and may even be faulty. For example, if an instructor has a small class, he may explain an event, then ask the class questions about it. If these reveal a lack of understanding, the instructor has only crude methods of discriminating what portion of his explanation was not adequate. If a student understands due to variables *other* than the instructor's explanation, then inadequate techniques may get strengthened. (I suspect that many of you will agree that bright students play this trick on us often.) Under the worst conditions, e.g., in the case of the college professor, weeks may elapse before he gets real feedback in the form of responses to examination questions. As many of us know, this feedback may turn out to be of the most punishing sort. One cannot even derive comfort from the students who give satisfactory answers; they may have gotten an understanding of the subject matter in spite of our efforts rather than because of them. According to this analysis, then, excellence in teaching is difficult to come by because the appropriate conditions for the shaping of teaching excellence are not met in the classroom situation; differential reinforcement is meager and the contingencies imprecise.

Consider, on the other hand, the machine teacher or programmer, the person who constructs the material for learning by successive approximations. Every phrase literally, gets differential reinforcement; if students consistently give an incorrect answer to a particular frame, that frame or the ones preceding it must be revised or supplemented by additional frames to bring the desired behavior up to strength. There is no opportunity in this venture to berate the student for "inattention" or "laziness."

He has answered the preceding frames; he has executed the behavior which was intended to prepare him for what followed. The responsibility for the subject's behavior falls directly on the teacher or programmer. This is the ultimate consequence of the behavioristic position and seems certain to produce widespread and perhaps surprising changes in teaching techniques. It seems perfectly reasonable to say that the machine-teaching enterprise cannot fail, because of its self-correcting features. Never before has the teacher's behavior and the student's reactions to it been made public in such detail.

I am perhaps more sorry than you are that I am unable to present definitive data at this time on how well children learn arithmetic by this method. There are enough *related* data, however, to make it risky to the point of presumption to assume that self-teaching by successive approximations will not work. Ferster and Sapon (1958) have taught elementary German to adults by this method. Douglas Porter has taught spelling to children, and I have taught a portion of elementary college physics preceded by some requisite mathematics. None of us have encountered insurmountable difficulties. I do not wish to give the impression that we know very much about this method of teaching at this time, but I do wish to emphasize that we know enough . . . [to be] more confident than ever that self-teaching by successive approximations presents no problems which researchers cannot overcome. In fact, I will go so far as to predict that classrooms of the future, their walls lined with exotic machines, will resemble nothing so much as the emporiums of Las Vegas. I am even willing to bet that the players will be equally intense in their pursuit of reinforcements.

Teaching Machines

B. F. SKINNER

THERE ARE more people in the world than ever before, and a far greater part of them want an education. The demand cannot be met simply by building more schools and training more teachers. Education must become more efficient. To this end, curricula must be revised and simplified, and textbooks and classroom techniques improved. In any other field a demand for increased production would have led at once to the invention of laborsaving capital equipment. Education has reached this stage very late, possibly through a misconception of its task. Thanks to the advent of television, however, the so-called audio-visual aids are being re-examined. Film projectors, television sets, phonographs, and tape recorders are finding their way into American schools and colleges.

Audio-visual aids supplement and may even supplant lectures, demonstrations, and textbooks. In doing so, they serve one function of the teacher: They present material to the student and, when successful, make it so clear and interesting that the student learns. There is another function to which they contribute little or nothing. It is best seen in the productive interchange between teacher and student in the small classroom or tutorial situation. Much of that interchange has already been sacrificed in American education in order to teach large numbers of students. There is a real danger that it will be wholly obscured if use of equipment designed simply to *present* material becomes widespread. The student is becoming more and more a mere passive receiver of instruction.

PRESSEY'S TEACHING MACHINES

There is another kind of capital equipment which will encourage the student to take an active role in the instructional

---

Dr. Skinner's article is reprinted from *Science*, Vol. 128, October 24, 1958, with minor editorial changes in format. It also comprised Part I of the report by Skinner and Holland (1958) to the Fund for the Advancement of Education. (The remaining portions of this report appear as the next paper in the present volume.) Figure III of the original article, which showed several booths containing machines similar to the one shown here in Figure II, is not included in this volume.

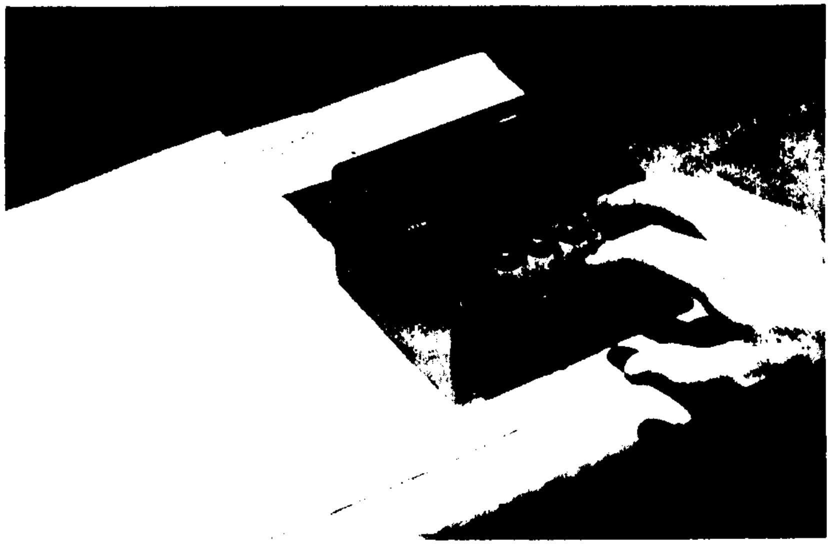


FIG. I. A recent model of Pressey's self-testing machine

process. The possibility was recognized in the 1920's, when Sidney L. Pressey designed several machines for the automatic testing of intelligence and information. A recent model of one of these is shown in Figure I. In using the device, the student refers to a numbered item in a multiple-choice test. He presses the button corresponding to his first choice of answer. If he is right, the device moves on to the next item; if he is wrong, the error is tallied, and he must continue to make choices until he is right.[1] Such machines, Pressey (1926) pointed out, could not only test and score, they could *teach*. When an examination is corrected and returned after a delay of many hours or days, the student's behavior is not appreciably modified. The immediate report supplied by a self-scoring device, however, can have an important instructional effect. Pressey also pointed out that such machines would increase efficiency in another way. Even in a small classroom the teacher usually knows that he is moving too slowly for some students and too fast for others. Those who could go faster are penalized, and those who should go slower are poorly taught and unnecessarily punished by criticism and failure. Machine instruction would permit each student to proceed at his own rate.

---

[1] The Navy's Self-Rater is a larger version of Pressey's machine. The items are printed on code-punched plastic cards fed by the machine. The time required to answer is taken into account in scoring.

The "industrial revolution in education" which Pressey envisioned stubbornly refused to come about. In 1932 he expressed his disappointment. "The problems of invention are relatively simple," he wrote. "With a little money and engineering resource, a great deal could easily be done. The writer has found from bitter experience that one person alone can accomplish relatively little and he is regretfully dropping further work on these problems. But he hopes that enough may have been done to stimulate other workers, that this fascinating field may be developed." (Pressey, 1932)

Pressey's machines succumbed in part to cultural inertia; the world of education was not ready for them. But they also had limitations which probably contributed to their failure. Pressey was working against a background of psychological theory which had not come to grips with the learning process. The study of human learnings was dominated by the "memory drum" and similar devices originally designed to study forgetting. Rate of learning was observed, but little was done to change it. Why the subject of such an experiment bothered to learn at all was of little interest. "Frequency" and "recency" theories of learning and principles of "massed and spaced practice" concerned the conditions under which responses were remembered.

Pressey's machines were designed against this theoretical background. As versions of the memory drum, they were primarily testing devices. They were to be used after some amount of learning had already taken place elsewhere. By confirming correct responses and by weakening responses which should not have been acquired, a self-testing machine does, indeed, teach; but it is not designed primarily for that purpose. Nevertheless, Pressey seems to have been the first to emphasize the importance of immediate feedback in education and to propose a system in which each student could move at his own pace. He saw the need for capital equipment in realizing these objectives. Above all, he conceived of a machine which (in contrast with the audio-visual aids which were beginning to be developed) permitted the student to play an active role.

X

ANOTHER KIND OF MACHINE

The learning process is now much better understood. Much of what we know has come from studying the behavior of lower organisms, but the results hold surprisingly well for human subjects. The emphasis in this research has not been on proving or disproving theories but on discovering and controlling the variables of which learning is a function. This practical orientation has paid off, for a surprising degree of control has been achieved. By arranging appropriate "contingencies of reinforcement," specific forms of behavior can be set up and brought under the control of specific classes of stimuli. The resulting behavior can be maintained in strength for long periods of time. A technology based on this work has already been put to use in neurology, pharmacology, nutrition, psychophysics, psychiatry, and elsewhere (Skinner, 1957a).

The analysis is also relevant to education. A student is "taught," in the sense that he is induced to engage in new forms of behavior and in specific forms upon specific occasions. It is not merely a matter of teaching him *what* to do; we are as much concerned with the probability that appropriate behavior will, indeed, appear at the proper time—an issue which would be classed traditionally under motivation. In education the behavior to be shaped and maintained is usually verbal, and it is to be brought under the control of both verbal and nonverbal stimuli. Fortunately, the special problems raised by verbal behavior can be submitted to a similar analysis (Skinner, 1957e).

If our current knowledge of the acquisition and maintenance of verbal behavior is to be applied to education, some sort of teaching machine is needed. Contingencies of reinforcement which change the behavior of lower organisms often cannot be arranged by hand; rather elaborate apparatus is needed. The human organism requires even more subtle instrumentation. An appropriate teaching machine will have several important features. The student must *compose* his response rather than select it from a set of alternatives, as in a multiple-choice self-rater. One reason for this is that we want him to recall rather than recognize—to make a response as well as see that it is right. Another reason is that effective multiple-choice material must contain plausible wrong responses, which are out of place in the delicate process of "shaping" behavior because they strengthen unwanted forms. Although it is much easier to build a machine to score multiple-choice answers than to evaluate a composed response, the technical advantage is outweighed by these and other considerations.

A second requirement of a minimal teaching machine also distinguishes it from earlier versions. In acquiring complex behavior, the student must pass through a carefully designed sequence of steps, often of considerable length. Each step must be so small that it can always be taken, yet in taking it the student moves somewhat closer to fully competent behavior. The machine must make sure that these steps are taken in a carefully prescribed order.

Several machines with the required characteristics have been built and tested. Sets of separate presentations or "frames" of visual material are stored on disks, cards, or tapes. One frame is presented at a time, adjacent frames being out of sight. In one type of machine the student composes a response by moving printed figures or letters (Skinner, 1954). His setting is compared by the machine with a coded response. If the two correspond, the machine automatically presents the next frame. If they do not, the response is cleared and another must be composed. The student cannot proceed to a second step until the first has been taken. A machine of this kind is being tested in teaching spelling, arithmetic, and other subjects in the lower grades.

For more advanced students—from junior high school, say, through college—a machine which senses an arrangement of letters or figures is unnecessarily rigid in specifying form of response. Fortunately, such students may be asked to compare their responses with printed material revealed by the machine. In the machine shown in Figure II, material is printed in 30 radial frames on a 12-inch disk. The student inserts the disk and closes the machine. He cannot proceed until the machine has been locked, and, once he has begun, the machine cannot be unlocked. All but a corner of one frame is visible through a window. The student writes his response on a paper strip exposed through a second opening. By lifting a lever on the front of the machine, he moves what he has written under a transparent cover and uncovers the correct response in the remaining corner of the frame. If the two responses correspond, he moves the lever horizontally. This movement punches a hole in the paper oppo-

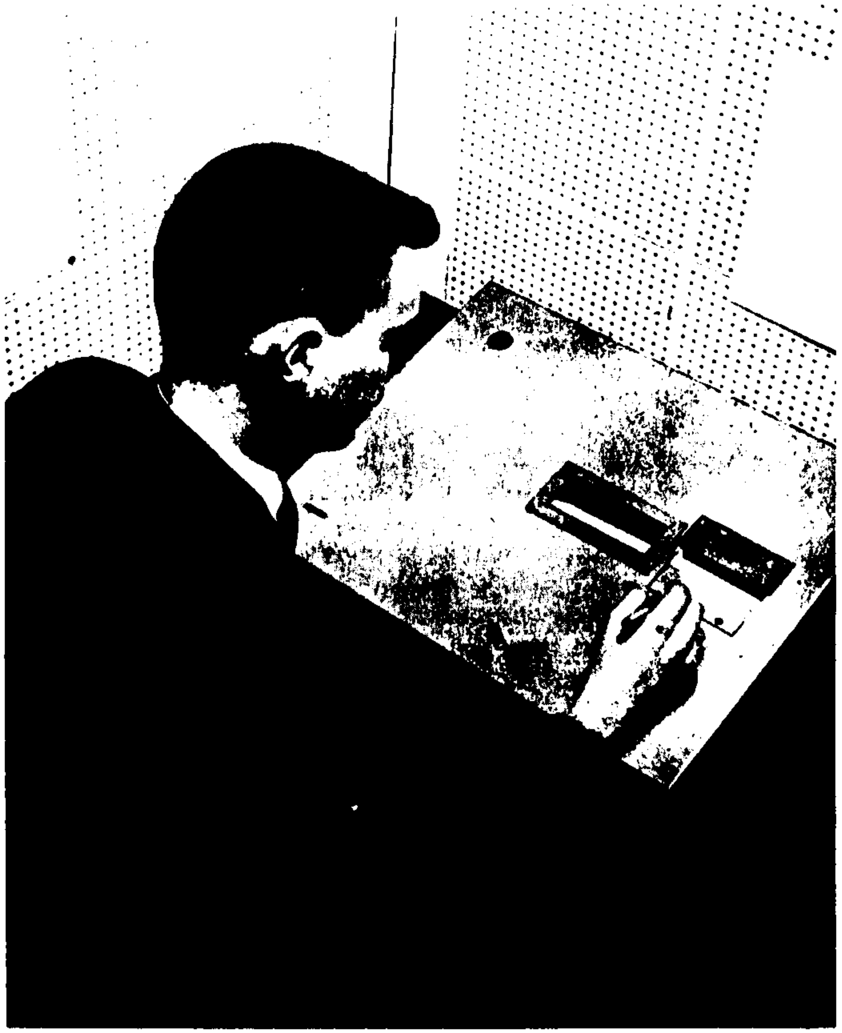


FIG. 11. Student at work in the self-instruction room. Material appears in the left-hand window. The student writes his response on a strip of paper exposed at the right.

site his response, recording the fact that he called it correct, and alters the machine so that the frame will not appear again when the student works around the disk a second time. Whether the response was correct or not, a second frame appears when

the lever is returned to its starting position. The student proceeds in this way until he has responded to all frames. He then works around the disk a second time, but only those frames appear to which he has not correctly responded. When the disk revolves without stopping, the assignment is finished. (The student is asked to repeat each frame until a correct response is made to allow for the fact that, in telling him that a response is wrong, such a machine tells him what is right.)

The machine itself, of course, does not teach. It simply brings the student into contact with the person who composed the material it presents. It is a laborsaving device because it can bring one programmer into contact with an indefinite number of students. This may suggest mass production, but the effect upon each student is surprisingly like that of a private tutor. The comparison holds in several respects: (a) There is a constant interchange between program and student. Unlike lectures, textbooks, and the usual audio-visual aids, the machine induces sustained activity. The student is always alert and busy. (b) Like a good tutor the machine insists that a given point be thoroughly understood, either frame by frame or set by set, before the student moves on. Lectures, textbooks, and their mechanized equivalents, on the other hand, proceed without making sure that the student understands and easily leave him behind. (c) Like a good tutor the machine presents just that material for which the student is ready. It asks him to take only that step which he is at the moment best equipped and most likely to take. (d) Like a skillful tutor the machine helps the student to come up with the right answer. It does this in part through the orderly construction of the program and in part with techniques of hinting, prompting, suggesting, and so on, derived from an analysis of verbal behavior (Skinner, 1957e). (e) Lastly, of course, the machine, like the private tutor, reinforces the student for every correct response, using this immediate feedback not only to shape his behavior most efficiently but to maintain it in strength in a manner which the layman would describe as "holding the student's interest."

PROGRAMMING MATERIAL

The success of such a machine depends on the material used in it. The task of programming a given subject is at first sight rather formidable. Many helpful techniques can be derived from a general analysis of the relevant behavioral processes, verbal and nonverbal. Specific forms of behavior are to be evoked and, through differential reinforcement, brought under the control of specific stimuli.

This is not the place for a systematic review of available techniques or of the kind of research which may be expected to discover others. However, the machines themselves cannot be adequately described without giving a few examples of programs. We may begin with a set of frames (see Table 1) designed to teach a third- or fourth-grade pupil to spell the word *manufacture*. The six frames are presented in the order shown, and the pupil moves sliders to expose letters in the open squares.

Table 1. A SET OF FRAMES DESIGNED TO TEACH A THIRD- OR FOURTH-GRADE PUPIL TO SPELL THE WORD "MANUFACTURE"

1. **Manufacture** means to make or build. *Chair factories manufacture chairs.* Copy the word here:

   □ □ □ □ □ □ □ □ □ □ □

2. Part of the word is like part of the word **factory**. Both parts come from an old word meaning *make* or *build.*

   m a n u □ □ □ □ u r e

3. Part of the word is like part of the word **manual**. Both parts come from an old word for *hand*. Many things used to be made by hand.

   □ □ □ □ f a c t u r e

4. The same letter goes in both spaces:

   m □ n u f □ c t u r e

5. The same letter goes in both spaces:

   m a n □ f a c t □ r e

6. Chair factories □ □ □ □ □ □ □ □ □ □ □ chairs.

The word to be learned appears in bold face in Frame 1, with an example and a simple definition. The pupil's first task is simply to copy it. When he does so correctly, Frame 2 appears. He must now copy selectively: he must identify "fact" as the common part of "manufacture" and "factory." This helps him to spell the word and also to acquire a separable "atomic" verbal operant (Skinner, 1957e). In Frame 3 another root must be copied selectively from "manual." In Frame 4 the pupil must for the first time insert letters without copying. Since he is asked to insert the same letter in two places, a wrong response will be doubly conspicuous, and the chance of failure is thereby minimized. The same principle governs Frame 5. In Frame 6 the pupil spells the word to complete the sentence used as an example in Frame 1. Even a poor student is likely to do this correctly because he has just composed or completed the word five times, has made two important root-responses, and has learned that two letters occur in the word twice. He has probably learned to spell the word without having made a mistake.

Teaching spelling is mainly a process of shaping complex forms of behavior. In other subjects—for example, arithmetic—responses must be brought under the control of appropriate stimuli. Unfortunately, the material which has been prepared for teaching arithmetic[2] does not lend itself to excerpting. The numbers 0 through 9 are generated in relation to objects, quantities, and scales. The operations of addition, subtraction, multiplication, and division are thoroughly developed before the number 10 is reached. In the course of this the pupil composes equations and expressions in a great variety of alternative forms. He completes not only $5 + 4 = \square$, but $\square + 4 = 9$, $5 \square 4 = 9$, and so on, aided in most cases by illustrative materials. No appeal is made to rote memorizing, even in the later acquisition of the tables. The student is expected to arrive at $9 \times 7 = 63$, not by memorizing it as he would memorize a line of poetry, but by putting into practice such principles as that 9 times a number is the same as 10 times the number minus the number (both of these being "obvious" or already well learned), that the digits in a multiple of 9 add to 9, that in composing successive multiples of 9 one counts backwards (9, 18, 27, 36, and so on), that 9 times a single digit is a number beginning with one less than the digit (9 times 6 is 50 something), and possibly even that the product of two numbers separated by only one number is equal to the square of the separating number minus 1 (the square of 8 already being familiar from a special series of frames concerned with squares).

---

[2] This material was prepared with the assistance of Susan R. Meyer.

Programs of this sort run to great length. At five or six frames per word, four grades of spelling may require 20,000 or 25,000 frames; and three or four grades of arithmetic, as many again. If these figures seem large, it is only because we are thinking of the normal contact between teacher and pupil. Admittedly, a teacher cannot supervise 10,000 or 15,000 responses made by each pupil per year. But the pupil's time is not so limited. In any case, surprisingly little time is needed. Fifteen minutes per day on a machine should suffice for each of these programs, the machines being free for other students for the rest of each day. (It is probably because traditional methods are so inefficient that we have been led to suppose that education requires such a prodigious part of a young person's day.)

A simple technique used in programming material at the high-school or college level, by means of the machine shown in Figure II, is exemplified in teaching a student to recite a poem. The first line is presented with several unimportant letters omitted. The student must read the line "meaningfully" and supply the missing letters. The second, third, and fourth frames present succeeding lines in the same way. In the fifth frame, the first line reappears with other letters also missing. Since the student has recently read the line, he can complete it correctly. He does the same for the second, third, and fourth lines. Subsequent frames are increasingly incomplete, and eventually—say, after 20 or 24 frames—the student reproduces all four lines without external help and quite possibly without having made a wrong response. The technique is similar to that used in teaching spelling: Responses are first controlled by a text, but this is slowly reduced (colloquially, "vanished") until the responses can be emitted without a text, each member in a series of responses being now under the "intraverbal" control of other members.

"Vanishing" can be used in teaching other types of verbal behavior. When a student describes the geography of part of the world or the anatomy of part of the body, or names plants and animals from specimens or pictures, verbal responses are controlled by nonverbal stimuli. In setting up such behavior, the student is first asked to report features of a fully labeled map, picture, or object; and the labels are then vanished. In teaching a map, for example, the machine asks the student to describe spatial relations among cities, countries, rivers, and so on, as shown on a fully labeled map. He is then asked to do the same with a map in which the names are incomplete or, possibly, lacking. Eventually he is asked to report the same relations with no map at all. If the material has been well programmed, he can do so correctly. Instruction is sometimes concerned not so much with imparting a new repertoire of verbal responses as with getting the student to describe something accurately in any available terms. The machine can "make sure the student understands" a graph, diagram, chart, or picture by asking him to identify and explain its features—correcting him, of course, whenever he is wrong.

In addition to charts, maps, graphs, models, and so on, the student may have access to auditory material. In learning to take dictation in a foreign language, for example, he selects a short passage on an indexing phonograph according to instructions given by the machine. He listens to the passage as often as necessary and then transcribes it. The machine then reveals the correct text. The student may listen to the passage again to discover the sources of any error. The indexing phonograph may also be used with the machine to teach other language skills, as well as telegraphic code, music, speech, parts of literary and dramatic appreciation, and other subjects.

A typical program combines many of these functions. The set of frames shown in Table 2 is designed to induce the student of high-school physics to talk intelligently, and to some extent technically, about the emission of light from an incandescent source. In using the machine, the student will write a word or phrase to complete a given item and then uncover the corresponding word or phrase shown here in the column at the right. The reader who wishes to get the "feel" of the material should cover the right-hand column with a card, uncovering each line only after he has completed the corresponding item.

Several programming techniques are exemplified by the set of frames in Table 2. Technical terms are introduced slowly. For example, the familiar term "fine wire" in Frame 2 is followed by a definition of the technical term "filament" in Frame 4; "filament" is then asked for in the presence of the nonscientific synonym in Frame 5 and without the synonym in Frame 9. In the same way, "glow," "give off light," and "send out light" in early frames are followed by a definition of "emit" with a synonym in Frame 7. Various inflected forms of "emit" then

Table 2. PART OF A PROGRAM IN HIGH-SCHOOL PHYSICS

The machine presents one item at a time. The student completes the item and then uncovers the corresponding word or phrase shown at the right.

SENTENCE TO BE COMPLETED	WORD TO BE SUPPLIED
1. The important parts of a flashlight are the battery and the bulb. When we "turn on" a flashlight, we close a switch which connects the battery with the ______.	bulb
2. When we turn on a flashlight, an electric current flows through the fine wire in the ______ and causes it to grow hot.	bulb
3. When the hot wire glows brightly, we say that it gives off or sends out heat and ______.	light
4. The fine wire in the bulb is called a filament. The bulb "lights up" when the filament is heated by the passage of a(n) ______ current.	electric
5. When a weak battery produces little current, the fine wire, or ______, does not get very hot.	filament
6. A filament which is less hot sends out or gives off ______ light.	less
7. "Emit" means "send out." The amount of light sent out, or "emitted," by a filament depends on how ______ the filament is.	hot
8. The higher the temperature of the filament the ______ the light emitted by it.	brighter, stronger
9. If a flashlight battery is weak, the ______ in the bulb may still glow, but with only a dull red color.	filament
10. The light from a very hot filament is colored yellow or white. The light from a filament which is not very hot is colored ______.	red
11. A blacksmith or other metal worker sometimes makes sure that a bar of iron is heated to a "cherry red" before hammering it into shape. He uses the ______ of the light emitted by the bar to tell how hot it is.	color
12. Both the color and the amount of light depend on the ______ of the emitting filament or bar.	temperature
13. An object which emits light because it is hot is called "incandescent." A flashlight bulb is an incandescent source of ______.	light
14. A neon tube emits light but remains cool. It is, therefore, not an incandescent ______ of light.	source
15. A candle flame is hot. It is a(n) ______ source of light.	incandescent
16. The hot wick of a candle gives off small pieces or particles of carbon which burn in the flame. Before or while burning, the hot particles send out, or ______, light.	emit
17. A long candlewick produces a flame in which oxygen does not reach all the carbon particles. Without oxygen the particles cannot burn. Particles which do not burn rise above the flame as ______.	smoke
18. We can show that there are particles of carbon in a candle flame, even when it is not smoking, by holding a piece of metal in the flame. The metal cools some of the particles before they burn, and the unburned carbon ______ collect on the metal as soot.	particles
19. The particles of carbon in soot or smoke no longer emit light because they are ______ than when they were in the flame.	cooler, colder
20. The reddish part of a candle flame has the same color as the filament in a flashlight with a weak battery. We might guess that the yellow or white parts of a candle flame are ______ than the reddish part.	hotter
21. "Putting out" an incandescent electric light means turning off the current so that the filament grows too ______ to emit light.	cold, cool
22. Setting fire to the wick of an oil lamp is called ______ the lamp.	lighting
23. The sun is our principal ______ of light, as well as of heat.	source
24. The sun is not only very bright but very hot. It is a powerful ______ source of light.	incandescent
25. Light is a form of energy. In "emitting light" an object changes, or "converts," one form of ______ into another.	energy
26. The electrical energy supplied by the battery in a flashlight is converted to ______ and ______.	heat, light; light, heat
27. If we leave a flashlight on, all the energy stored in the battery will finally be changed or ______ into heat and light.	converted
28. The light from a candle flame comes from the ______ released by chemical changes as the candle burns.	energy
29. A nearly "dead" battery may make a flashlight bulb warm to the touch, but the filament may still not be hot enough to emit light—in other words, the filament will not be ______ at that temperature.	incandescent
30. Objects, such as a filament, carbon particles, or iron bars, become incandescent when heated to about 800 degrees Celsius. At that temperature they begin to ______ ______.	emit light
31. When raised to any temperature above 800 degrees Celsius, an object such as an iron bar will emit light. Although the bar may melt or vaporize, its particles will be ______ no matter how hot they get.	incandescent
32. About 800 degrees Celsius is the lower limit of the temperature at which particles emit light. There is no upper limit of the ______ at which emission of light occurs.	temperature
33. Sunlight is ______ by very hot gases near the surface of the sun.	emitted
34. Complex changes similar to an atomic explosion generate the great heat which explains the ______ of light by the sun.	emission
35. Below about ______ degrees Celsius an object is not an incandescent source of light.	800

follow, and "emit" itself is asked for with a synonym in Frame 16. It is asked for without a synonym, but in a helpful phrase, in Frame 30, and "emitted" and "emission" are asked for without help in Frames 33 and 34. The relation between temperature and amount and color of light is developed in several frames before a formal statement using the word "temperature" is asked for in Frame 12. "Incandescent" is defined and used in Frame 13, is used again in Frame 14, and is asked for in Frame 15—the student receiving a thematic prompt from the recurring phrase "incandescent source of light." A formal prompt is supplied by "candle." In Frame 25 the new response, "energy," is easily evoked by the words "form of ———" because the expression "form of energy" is used earlier in the frame. "Energy" appears in the next two frames and is finally asked for, without aid, in Frame 28. Frames 30 through 35 discuss the limiting temperatures of incandescent objects, while reviewing several kinds of sources. The figure 800 is used in three frames. Two intervening frames then permit some time to pass before the response "800" is asked for.

Unwanted responses are eliminated with special techniques. If, for example, the second sentence in Frame 24 were simply "It is a(n) ——— source of light," the two "very's" would frequently lead the student to fill the blank with "strong" or a synonym thereof. This is prevented by inserting the word "powerful" to make a synonym redundant. Similarly, in Frame 3 the words "heat and" pre-empt the response "heat," which would otherwise correctly fill the blank.

The net effect of such material is more than the acquisition of facts and terms. Beginning with a largely unverbalized acquaintance with flashlights, candles, and so on, the student is induced to talk about familiar events, together with a few new facts, with a fairly technical vocabulary. He applies the same terms to facts which he may never before have seen to be similar. The emission of light from an incandescent source takes shape as a topic or field of inquiry. An understanding of the subject emerges which is often quite surprising in view of the fragmentation required in item-building.

It is not easy to construct such a program. Where a confusing or elliptical passage in a textbook is forgivable because it can be clarified by the teacher, machine material must be self-contained and wholly adequate. There are other reasons why textbooks, lecture outlines, and film scripts are of little help in preparing a program. They are usually not logical or developmental arrangements of material but strategems which the authors have found successful under existing classroom conditions. The examples they give are more often chosen to hold the student's interest than to clarify terms and principles. In composing material for the machine, the programmer may go directly to the point.

A first step is to define the field. A second is to collect technical terms, facts, laws, principles, and cases. These must then be arranged in a plausible developmental order—linear if possible, branching if necessary. A mechanical arrangement, such as a card-filing system, helps. The material is distributed among the frames of a program to achieve an arbitrary density. In the final composition of an item, techniques for strengthening asked-for responses and for transferring control from one variable to another are chosen from a list, according to a given schedule, in order to prevent the establishment of irrelevant verbal tendencies appropriate to a single technique. When one set of frames has been composed, its terms and facts are seeded mechanically among succeeding sets, where they will be referred to in composing later items to make sure that the earlier repertoire remains active. Thus, the technical terms, facts, and examples in Table 2 have been distributed for re-use in succeeding sets on reflection, absorption, and transmission, where they are incorporated into items dealing mainly with other matters. Sets of frames for explicit review can, of course, be constructed. Further research will presumably discover other, possibly more effective, techniques. Meanwhile, it must be admitted that a considerable measure of art is needed in composing a successful program.

Whether good programming is to remain an art or to become a scientific technology, it is reassuring to know that there is a final authority—the student. An unexpected advantage of machine instruction has proved to be the feedback to the *programmer*. In the elementary-school machine, provision is made for discovering which frames commonly yield wrong responses; and in the high-school and college machine, the paper strips bearing written answers are available for analysis. A trial run of the first version of a program quickly reveals frames which need to be altered or sequences which need to be lengthened.

One or two revisions in the light of a few dozen responses work a great improvement. No comparable feedback is available to the lecturer, textbook writer, or maker of films. Although one text or film may seem to be better than another, it is usually impossible to say, for example, that a given sentence on a given page or a particular sequence in a film is causing trouble.

Difficult as programming is, it has its compensations. It is a salutary thing to try to guarantee a right response at every step in the presentation of a subject matter. The programmer will usually find that he has been accustomed to leave much to the student—that he has frequently omitted essential steps and neglected to invoke relevant points. The responses made to his material may reveal surprising ambiguities. Unless he is lucky, he may find that he still has something to learn about his subject. He will almost certainly find that he needs to learn a great deal more about the behavioral changes he is trying to induce in the student. This effect of the machine in confronting the programmer with the full scope of his task may in itself produce a considerable improvement in education.

Composing a set of frames can be an exciting exercise in the analysis of knowledge. The enterprise has obvious bearings on scientific methodology. There are hopeful signs that the epistemological implications will induce experts to help in composing programs. The expert may be interested for another reason. We can scarcely ask a topflight mathematician to write a primer in second-grade arithmetic if it is to be used by the average teacher in the average classroom. But a carefully controlled machine presentation and the resulting immediacy of contact between programmer and student offer a very different prospect, which may be enough to induce those who know most about the subject to give some thought to the nature of arithmetical behavior and to the various forms in which such behavior should be set up and tested.

CAN MATERIAL BE TOO EASY?

The traditional teacher may view these programs with concern. He may be particularly alarmed by the effort to maximize success and minimize failure. He has found that students do not pay attention unless they are worried about the consequences of their work. The customary procedure has been to maintain the necessary anxiety by inducing errors. In recitation, the student who obviously knows the answer is not too often asked; a test item which is correctly answered by everyone is discarded as nondiscriminating; problems at the end of a section in a textbook in mathematics generally include one or two very difficult items; and so on. (The teacher-turned-programmer may be surprised to find this attitude affecting the construction of items. For example, he may find it difficult to allow an item to stand which "gives the point away." Yet if we can solve the motivational problem with other means, what is more effective than giving a point away?) Making sure that the student knows he doesn't know is a technique concerned with motivation, not with the learning process. Machines solve the problem of motivation in other ways. There is no evidence that what is easily learned is more readily forgotten. If this should prove to be the case, retention may be guaranteed by subsequent material constructed for an equally painless review.

The standard defense of "hard" material is that we want to teach more than subject matter. The student is to be challenged and taught to "think." The argument is sometimes little more than a rationalization for a confusing presentation, but it is doubtless true that lectures and texts are often inadequate and misleading by design. But to what end? What sort of "thinking" does the student learn in struggling through difficult material? It is true that those who learn under difficult conditions are better students. But are they better because they have surmounted difficulties, or do they surmount them because they are better? In the guise of teaching thinking we set difficult and confusing situations and claim credit for the students who deal with them successfully.

The trouble with deliberately making education difficult in order to teach thinking is (a) that we must remain content with the students thus selected, even though we know that they are only a small part of the potential supply of thinkers and (b) that we must continue to sacrifice the teaching of subject matter by renouncing effective but "easier" methods. A more sensible program is to analyze the behavior called "thinking" and produce it according to specifications. A program specifically concerned with such behavior could be composed of material already available in logic, mathematics, scientific method, and psychology. Much would doubtless be added in completing an effective program. The machine has already yielded important relevant by-products. Immediate feedback encourages a more careful reading of programmed material than is the case in studying a text, where the consequences of attention or inattention are so long deferred that they have little effect on reading skills. The behavior involved in observing or attending to detail—as in inspecting charts and models or listening closely to recorded speech —is efficiently shaped by the contingencies arranged by the machine. And when an immediate result is in the balance, a student will be more likely to learn how to marshal relevant material, to concentrate on specific features of a presentation, to reject irrelevant materials, to refuse the easy but wrong solution, and to tolerate indecision, all of which are involved in effective thinking.

Part of the objection to easy material is that the student will come to depend on the machine and will be less able than ever to cope with the inefficient presentations of lectures, textbooks, films, and "real life." This is indeed a problem. All good teachers must "wean" their students, and the machine is no exception. The better the teacher, the more explicit must the weaning process be. The final stages of a program must be so designed that the student no longer requires the helpful conditions arranged by the machine. This can be done in many ways—among others by using the machine to discuss material which has been studied in other forms. These are questions which can be adequately answered only by further research.

No large-scale "evaluation" of machine teaching has yet been attempted. We have so far been concerned mainly with practical problems in the design and use of machines and with testing and revising sample programs. The machine shown in Figure II was built and tested with a grant from the Fund for the Advancement of Education. Material has been prepared and tested with the collaboration of Lloyd E. Homme, Susan R. Meyer, and James G. Holland.[3] A self-instruction room . . .

was set up under this grant. It contains 10 machines and was recently used to teach part of a course in human behavior to Harvard and Radcliffe undergraduates. Nearly 200 students completed 48 disks (about 1400 frames) prepared with the collaboration of Holland. The factual core of the course was covered, corresponding to about 200 pages of the text (Skinner, 1953a). The median time required to finish 48 disks was 14.5 hours. The students were not examined on the material but were responsible for the text which overlapped it. Their reactions to the material and to self-instruction in general have been studied through interviews and questionnaires. Both the machines and the material are now being modified in the light of this experience, and a more explicit evaluation will then be made.

Meanwhile, it can be said that the expected advantages of machine instruction were generously confirmed. Unsuspected possibilities were revealed which are now undergoing further exploration. Although it is less convenient to report to a self-instruction room than to pick up a textbook in one's room or elsewhere, most students felt that they had much to gain in studying by machine. Most of them worked for an hour or more with little effort, although they often felt tired afterwards, and they reported that they learned much more in less time and with less effort than in conventional ways. No attempt was made to point out the relevance of the material to crucial issues, personal or otherwise, but the students remained interested. (Indeed, one change in the reinforcing contingencies suggested by the experiment is intended to *reduce* the motivational level.) An important advantage proved to be that the student always knew where he stood, without waiting for an hour test or final examination.

SOME QUESTIONS

Several questions are commonly asked when teaching machines are discussed. Cannot the results of laboratory research on learning be used in education without machines? Of course, they can. They should lead to improvements in textbooks, films, and other teaching materials. Moreover, the teacher who really understands the conditions under which learning takes place will be more effective, not only in teaching subject matter but in managing the class. Nevertheless, some sort of device is necessary to arrange the subtle contingencies of reinforcement required for optimal learning if each student is to have individual attention. In nonverbal skills this is usually obvious; texts and instructor can guide the learner but they cannot arrange the final contingencies which set up skilled behavior. It is true that the verbal skills at issue here are especially dependent upon social reinforcement, but it must not be forgotten that the machine simply mediates an *essentially verbal relation*. In shaping and maintaining verbal knowledge, we are not committed to the contingencies arranged through immediate personal contact.

---

[3] Dr. Homme prepared sets of frames for teaching part of college physics (kinematics), and Mrs. Meyer has prepared and informally tested material in remedial reading and vocabulary building at the junior high-school level. Others who have contributed to the development of teaching machines should be mentioned. Nathan H. Azrin cooperated with me in testing a version of a machine to teach arithmetic. C. B. Ferster and Stanley M. Sapon (1958) used a simple "machine" to teach German. Douglas Porter (1958c), of the Graduate School of Education at Harvard, has made an independent schoolroom test of machine instruction in spelling. Devra Cooper has experimented with the teaching of English composition for freshmen at the University of Kentucky. Thomas F. Gilbert, of the University of Georgia, has compared standard and machine instruction in an introductory course in psychology and, with the collaboration of J. E. Jewett, has prepared material in algebra. The U. S. Naval Training Devices Center has recently contracted with the University of Pennsylvania for a study of programs relating to the machine instruction of servicemen, under the direction of Eugene H. Galanter.

Machines may still seem unnecessarily complex compared with other mediators such as workbooks or self-scoring test forms. Unfortunately, these alternatives are not acceptable. When material is adequately programmed, adjacent steps are often so similar that one frame reveals the response to another. Only some sort of mechanical presentation will make successive frames independent of each other. Moreover, in self-instruction an automatic record of the student's behavior is especially desirable, and for many purposes it should be foolproof. Simplified versions of the present machines have been found useful—for example, in the work of Ferster and Sapon (1958) or Porter (1958c), and of Gilbert[4]—but the mechanical and economic problems are so easily solved that a machine with greater capabilities is fully warranted.

Will machines replace teachers? On the contrary, they are capital equipment to be used by teachers to save time and labor. In assigning certain mechanizable functions to machines, the teacher emerges in his proper role as an indispensable human being. He may teach more students than heretofore—this is probably inevitable if the world-wide demand for education is to be satisfied—but he will do so in fewer hours and with fewer burdensome chores. In return for his greater productivity he can ask society to improve his economic condition.

---

[4] See Footnote 3, p. 154-55.

The role of the teacher may well be changed, for machine instruction will affect several traditional practices. Students may continue to be grouped in "grades" or "classes," but it will be possible for each to proceed at his own level, advancing as rapidly as he can. The other kind of "grade" will also change its meaning. In traditional practice a C means that a student has a smattering of a whole course. But if machine instruction assures mastery at every stage, a grade will be useful only in showing *how far* a student has gone; C might mean that he is halfway through a course. Given enough time he will be able to get an A; and since A is no longer a motivating device, this is fair enough. The quick student will meanwhile have picked up A's in other subjects.

Differences in ability raise other questions. A program designed for the slowest student in the school system will probably not seriously delay the fast student, who will be free to progress at his own speed. (He may profit from the full coverage by filling in unsuspected gaps in his repertoire.) If this does not prove to be the case, programs can be constructed at two or more levels, and students can be shifted from one to the other as performances dictate. If there are also differences in "types of thinking," the extra time available for machine instruction may be used to present a subject in ways appropriate to many types. Each student will presumably retain and use those ways which he finds most useful. The kind of individual difference which arises simply because a student has missed part of an essential sequence (compare the child who has no "mathematical ability" because he was out with the measles when fractions were first taken up) will simply be eliminated.

OTHER USES

Self-instruction by machine has many special advantages apart from educational institutions. Home study is an obvious case. In industrial and military training it is often inconvenient to schedule students in groups, and individual instruction by machine should be a feasible alternative. Programs can also be constructed in subjects for which teachers are not available—for example, when new kinds of equipment must be explained to operators and repairmen or where a sweeping change in method finds teachers unprepared (Menger, 1958). Education sometimes fails because students have handicaps which make a normal relationship with a teacher difficult or impossible. (Many blind children are treated today as feeble-minded because no one has had the time or patience to make contact with them. Deaf-mutes, spastics, and others suffer similar handicaps.) A teaching machine can be adapted to special kinds of communication—for example, Braille—and, above all, it has infinite patience.

CONCLUSION

An analysis of education within the framework of a science of behavior has broad implications. Our schools, in particular our "progressive" schools, are often held responsible for many current problems—including juvenile delinquency and the threat of a more powerful foreign technology. One remedy frequently suggested is a return to older techniques, especially to a greater "discipline" in schools. Presumably this is to be obtained with some form of punishment, to be administered either with certain classical instruments of physical injury—the dried bullock's tail of the Greek teacher or the cane of the English schoolmaster—or as disapproval or failure, the frequency of which is to be increased by "raising standards." This is probably not a feasible solution. Not only education but Western culture as a whole is moving away from aversive practices. We cannot prepare young people for one kind of life in institutions organized on quite different principles. The discipline of the birch rod may facilitate learning, but we must remember that it also breeds followers of dictators and revolutionists.

In the light of our present knowledge a school system must be called a failure if it cannot induce students to learn except by threatening them for not learning. That this has always been the standard pattern simply emphasizes the importance of modern techniques. John Dewey was speaking for his culture and his time when he attacked aversive educational practices and appealed to teachers to turn to positive and humane methods. What he threw out should have been thrown out. Unfortunately, he had too little to put in its place. Progressive education has been a temporizing measure which can now be effectively supplemented. Aversive practices can not only be replaced, they can be replaced with far more powerful techniques. The possibilities should be thoroughly explored if we are to build an educational system which will meet the present demand without sacrificing democratic principles.

X

The Use of Teaching Machines In College Instruction

B. F. SKINNER and JAMES G. HOLLAND

PART II

PRINCIPAL FEATURES OF A USEFUL TEACHING MACHINE

OUR EXPERIENCE during the past two years has revealed the following useful characteristics of teaching machines at the college level. The following terms will be used. A *program* is a set of materials on a given subject. It consists of a number of *sets of frames*. Each *set* is printed on a *disk* (present machine) or *tape* (new model). *Tape* and *disk* thus refer to a *set* of frames. *Frame* always refers to the material presented at any one moment. An *item* is the material in a frame. A *panel* is a chart, graph, or passage of text accessible during work on a *set*.

Storage of Material

Storing material on disks has proved to have several serious disadvantages: The space available per frame is small and awkwardly shaped, and the length of the set is fixed. It is often difficult to write items short enough for the space available if there are a reasonable number of frames per disk and the disk is not unmanageably large. In improving a set of frames, we cannot increase the number available on a disk. Another disadvantage is that the student can read the disk before inserting it in the machine. It is true that he does so in a highly motivated condition, but the interval between reading and writing answers when the disk has been inserted in the machine may be too brief for effective learning.

A revised model of the machine will solve these problems by

This paper is a reproduction of three of the four parts of a report submitted by Harvard University to the Fund for the Advancement of Education, August 15, 1958. Part I of this report was substantially identical to the immediately preceding article by Dr. Skinner published in the October 1958 issue of *Science*.

Help them "acquire" desired decision-making / problem-solving behaviors so they will "voluntarily" choose what's best for the collective good. This way the individual has "ownership" of the process and "feels" the choice was his to make.

using tapes stored in cartridges. The cartridges cannot conveniently be unrolled outside the machine, so that the student will not be able to read the material in advance. Much more space will be available both for the item and for the response. Increased space for the item will mean that more material can be presented before a response is asked for and also that diagrams, etc., may be used in separate frames, rather than on panels. More space for the correct *response* will permit listing acceptable synonyms and explaining why other common responses are unacceptable.

Record of the Student's Responses

In the revised machine the strip of paper on which responses are written during the first cycle will be presented again during the second cycle, except that correct responses will be skipped. The student will thus be reminded of his earlier response, and of its being wrong, and can alter the response conveniently without fully rewriting it. Much more readily detected punches will be used, and they will indicate *wrong* responses. The response strips will therefore be more readily analyzed in improving the program. In order to prevent cheating, especially when the machine is used to test the student, the response strip will remain in the machine when the student has finished. It will thus be impossible to alter it, to cover up dishonest entries, or to throw away one strip and submit a second.

Mode of Operation

Some students, particularly in the early stages of the experiment, found that they tended to move the lever to the right, thus calling their response right, before comparing the revealed response. In the revised machine the student will set up the judgment "right" or "wrong" in the form which can be corrected before it is irrevocably recorded. The new machine will be symmetrical, so that it will be equally convenient for left-handed and right-handed students. Lighting will be improved by eliminating reflecting surfaces. The writing surface will be more convenient.

There was a strong tendency to put down a quick, careless response in order to uncover the correct response in the machine. The more highly motivated the student, the stronger the tendency appeared to be. Many students were content to get the response right the second time around. It seems to be necessary to make correct first-cycle responses more important. This can be done by adding a slight punishment for a wrong response —for example, there could be a five-second delay in moving on to the next frame when a response has been judged incorrect. Or we could appeal to the punitive techniques of educational institutions and keep a record of the number of first-cycle mistakes. The student could be asked to keep such a record himself to emphasize his level of achievement. (Since there is an indelible record of the student's behavior in calling responses right or wrong, the simplest thing would be to ask him to copy the number of errors from a resetting counter. However, a printout counter could be built into the machine to stamp the number of mistakes on the strip.) Another possibility would be to set up a "par" score for each tape and to require that the student who fails to complete the tape at or under par on the first cycle open the machine and start again —this operation supplying him with a fresh response strip and requiring him to answer *all* items again on the first cycle before proceeding to the second.

It appears that material learned by machine should be occasionally tested. There are several possibilities: (a) Specific disks may be designed as tests—the student knowing this in advance, the machine not permitting him to go through his response strip a second time. (b) Tapes may be selected at random as tests, the student's first-cycle performance being stored for grading and the student repeating the tape with a fresh answer strip for subsequent instructional purposes. The student might or might not know in advance that he was working on a test tape.

Relation to Materials Not on the Tapes

A panel or other material which is to be available to the student while operating the machine should be brief. It should be arranged in outline form or in any other form which facilitates quick reference (e.g., important terms may be underlined). More use might be made of panels to list specific technical terms to be used in responding. A complete vocabulary appropriate to a given field may be printed on a panel, no word not in such a list being accepted as a correct response. This is not a multiple-choice system; the intention is to avoid nontechnical or otherwise unacceptable terms not on the list. The technique would be used only when special vocabulary distinctions are insisted upon.

The student may be required to read part of a text before operating a machine but not while operating. The machine simply checks the student's ability to comprehend the passage he has read, just as it checks his understanding of a graph or sketch of an apparatus. Later the technique may be used to "wean" the student from the machine (see Part I [Skinner, 1958c]).

PART III

PROGRAMMING VERBAL KNOWLEDGE

Certain general principles have evolved relating to the construction of programs.

Physical Properties

LENGTH OF PROGRAM. The length of a program is determined in the first place by the material covered and by the number of hours the student has available. It is possible that such a rough measure as a "one-semester" course may not suffice. In adapting a program to high school or college requirements, the programmer will presumably have to accept available time limits. This will not specify number of sets or frames because materials differ enormously in the speed with which students go through them. The following considerations about length of sets therefore bear on the length of programs.

LENGTH OF SET. Possibly the most important consideration is the normal length of session in the school or college schedule. Presumably more than one set will be covered by a student in one session. As indicated in Part IV, in the college material used this year the median time required for a disk (a set of 29 frames) was about 18 minutes, but some disks were so difficult that nearly an hour was needed by slower students. The time the student can ordinarily work without fatigue is important and may have no close relation to normal length of classroom session. A set of about 30 frames may be about the

right length. Anything below this means that the student returns to an item very quickly after making an incorrect response and may be getting it right for unimportant reasons. Much longer sets call for more time than the slow student has available in a single session. Moreover, a greater reinforcement from achievement may follow, even for the good student, from completing several sets per session. However, all statements of this sort depend upon the kind of material. In a quick vocabulary review, for example, where reading and writing time will be short, a set containing a hundred items might require only a few minutes. "Lead-in" sets can often be done quickly. In the long run, only an experimental analysis of material in a natural class situation will determine suitable length for a given kind of material.

LENGTH OF FRAME. It is sometimes desirable to present a considerable amount of text before asking the student to respond, particularly where every frame must reinstate a given situation to make it independent of earlier frames. There is a danger in lengthy material that the student will skip to the blank and attempt to find the answer from the adjacent text. If material is long, care must be taken to make sure that terms and materials *throughout* the text are relevant to the right answer.

NUMBER OF BLANKS. Sometimes, in economizing on space, two or more items can be printed in a given frame and separate answers required. When two blanks are asked for in the same item, however, it must be remembered that this may weaken or conceal the syntactical structure. The same behavior can usually be evoked by two frames with one blank in each while the student is puzzling out the syntax of an item with more than one blank. Syntax is not involved, however, when a student is asked, for example, to give both of two synonymous forms, one of which is to be technical and the other popular. Summarizing material in a set of multiple-blank items may also permit the student to organize it more coherently. Indeed, when the subject has been well prepared, it is often desirable to eliminate some of the syntax by use of multiple-blanks to further refine the stimuli controlling the response.

Composing the Program

SPECIFICATIONS OF A COURSE. The programmer must know what verbal behavior the student is to have in his repertoire after completing the course and how precisely and extensively he is to talk about the field.

Knowledge previously acquired. The student is assumed to possess some verbal behavior in the area before he starts the course. This must be stated, and the programmer must not at any time appeal to material not included in the statement or not provided by earlier parts of the program (e.g., a program of a college course assumes only the vocabulary and language ability of an entering freshman).

Ordering the knowledge to be acquired. At each step the programmer must ask, "What behavior must the student have before he can take this step?" A sequence of steps forms a progression from the initially assumed knowledge up to the specified final repertoire. No step should be encountered before the student has mastered everything needed to take it.

Listing terms, etc. Before writing frames for each set, the programmer should make lists of (a) the terms to be covered, (b) the processes or principles, (c) a wide range of illustrative examples. Ideally, examples should differ from each other in all respects except that in which they illustrate a term or principle (e.g., a list of examples of reflexes should include reflexes whose stimuli cover all possible sense modalities and reaction mechanisms). The programmer should also note various syntactical possibilities, including inflected forms (*elicit*, is *elicited by*, *elicitation*, etc.) and items "saying the same thing" but with different terms asked for. *(In the pupillary reflex a bright light elicits constriction of the pupil. The stimulus is (bright light).* And, *In the pupillary reflex the bright light is the (stimulus.)*

After these lists are made, a fairly mechanical procedure can be followed in writing frames so that each concept and process will be used with each example, in as wide a range or syntactical arrangements as possible.

At this stage the programmer should also consider what concepts can be appropriated and reviewed from earlier materials. This can be done by "leading in" with relevant concepts or by contrasting concepts (e.g., in contrasting operant behavior with respondent behavior, a set may review respondent behavior).

Writing frames. With rare exceptions, frames are statements with words missing, the missing word being supplied by the student. Multiple choice is not used, in order to avoid strengthening alternative wrong answers. (Exceptions are cases where the student has been prepared for a fine distinction which is then required in a multiple-choice item.) It is important that each frame be intelligible by itself because it will often be presented without the preceding frames in later cycles. However, a given set may assume a universe of discourse within which frames can be to some extent elliptical.

Getting Out New Responses Without Errors

1. Echoic and Textual Behavior and Formal Prompts

(a) *Panels.* An effective way of evoking new terms is to use printed material which remains before the student while he works through a disk or tape. Early frames guide him through careful textual behavior. The distinction between this and reading a text for memorization is that comprehensive reading is not assumed but rather forced by the frames which carry him through all points on the panel. The panel should include nothing which is not treated by the frames, because it should be as short as possible. The labor of searching the panel should be kept to a minimum. Reference letters or numbers are useful for this purpose in referring the student to particular parts of the panel. Subsequent tapes or disks should cover the same concepts and principles without the panel. These should contain new frames rather than repetitions of previous frames.

(b) *Use of new words in a series of frames.* Correct responses in a short series can be made to depend on careful observation of a new word. In the last frame in the series the student must write the word.

(c) *Definitions and examples.* These generate sentences easily completed. In an experiment described in more detail in Part IV, 95 out of 95 students answered "reinforce" to the frame: "A technical word for reward is reinforcement. To reward an organism with food is to ..... it with food." To the next frame: "Technically speaking, a thirsty organism can be ..... with water," 87 out of 95 answered "reinforced," seven answered "rewarded," one [answered] "conditioned."

(d) *Explicit formal prompts.* A response of low strength can be made more probable by giving its beginning or ending or selected letters. Indicating the number of letters is not very effective as a prompt, but permits the student to discard wrong responses.

(e) *Rhyming as a formal prompt.* A technique suitable for the young: 9 times 7 and just *1* more is 8 times 8 or ____."

2. Thematic Prompts

(a) "*Lead-Ins.*" Common knowledge, or phrases with high association value for the desired word, can sometimes be used. To the frame: "A doctor taps your knee (petaller tendon) with a rubber hammer to test your ____," 174 out of 182 students said "reflexes." To the following frame: "If your reflexes are normal, you ____ to the tap on the knee with a knee jerk," 107 out of 182 said "respond" and 72 said "react"—both responses being accepted as correct by the programmer. These already known terms could then be preserved and supplemented as part of a technical repertoire.

(b) *Indicating categories, sublanguages, or syntax.* A technical or lay term may be specifically asked for. That a response is to be a color-word is indicated in the sentence "Roses are colored ____," although "Roses are ____" would collect responses like "flowers," "beautiful," etc. Whether a response is to be a verb, noun, adjective, or adverb can sometimes be determined by supplying endings, such as *-ly*, or *-ed*.

(c) *Useful category.* "Opposites" is an example. To the item: "Reinforcement which consists of presenting stimuli (e.g., food or water) is called positive reinforcement; reinforcement which consists of terminating stimuli (e.g., loud noises or painful stimuli) is called ____ reinforcement," 173 out of 180 students said "negative." The result would be further assured by underlining "presenting," "terminating," and "positive" and by adding a phrase like "on the other hand" after the semicolon.

(d) *High-association words or common phrases.* To the item: "A reward simply makes it more ____ that an animal will behave in the same way again," 32 out of 95 students said "probable," 43 "likely," and 18 "certain," all of which were acceptable, the lapse in grammar in the last case being forgiven.

3. Eliminating Undesired Alternative Responses

The item: "Coins are conditioned ____ reinforcers" will bring out the desired response, "generalized," more often than: "Coins are ____ reinforcers," which will get many instances of "conditioned." Another solution is to have the student supply both alternatives, frequently with some indication of category or sublanguage. Thus an item might conclude: . . . "is called ____ or, in technical terms, ____."

PART IV

USE OF THE MACHINES IN TEACHING A GENERAL EDUCATION COURSE IN HUMAN BEHAVIOR*

Natural Sciences 114 is a second-level general education course. It may be elected from a rather long list of courses as a distribution requirement in natural science. The course has been given for eight years, and the enrollment is usually about 115, including both Harvard and Radcliffe students. At the early meetings of the course the teaching machine was explained, and students were invited to try out certain preliminary disks in the self-instruction room. It was explained that work on the machines would not be submitted to test, although students would be examined on those parts of the text (Skinner, 1953a) which covered some of the same material. The prospect of studying by machine was apparently responsible for an increase in enrollment to 187. Sophomores were most heavily represented, although there were many juniors and seniors and a number of freshmen admitted on special petition. The group also included 12 graduate students, members of the Academic Year Institute at Harvard.

Some disks were available at the beginning of the course, and more were made available as time passed. [The students were asked to complete all the machine work by the end of the first two months.] Forty-eight disks were prepared, each usually carrying 29 frames of material. Among the subjects covered were reflexes and Pavlovian conditioning, operant conditioning, discrimination, motivation, emotion, techniques of personal control, and illustrative or demonstrational material, in which the students were required to analyze and comment upon experiments, illustrative episodes, and so on. This was the basic "scientific" part of the course. Lectures and outside reading dealt with broader implications. Although the machine material overlapped the text to some extent, there was very little in common between the frames and the lectures.

---

* Spring term, 1958.

It had been assumed that in order to give students convenient access to machines they would be asked to reserve machines in advance. However, this proved not to be necessary. The self-instruction room was conveniently located, and students had little trouble, except just before the deadline, in finding a machine free. The room was usually open from 9 to 5 and from 7 to 10 in the evening, with a few hours on Saturday. (Deadline rushing could be reduced by setting a series of deadlines for parts of the course throughout the term.)

Most of the students paced themselves well during the term, though some left the disks till very near the deadline. One student did 31 disks in a single continuous session; and another, the entire set of 48 in only two days. Preliminary examination of their response strips does not indicate any great difference in performance. We have no evidence at the moment of any difference in retention.

The time required to complete a disk depended upon the difficulty. There are also wide individual differences among students. The median time required to go once through the easiest disk (to complete the "first cycle") was eight minutes, with a range of five to 13 minutes. The median time for answering all items correctly (by repeating the material) was 10 minutes, with a range of five to 19 minutes.

The most difficult disk showed a median time for the first cycle of 20 minutes, with a range of 10 to 48 minutes. All items on the most difficult disk were completed in a median time of 28 minutes with a range from 17 to 80 minutes. Both of these were atypical. The over-all median time for a disk was 15 minutes for the first cycle and 18 minutes to completion.

A typical student completed all 48 disks in about 14.5 hours. The fastest student used considerably less and the slow student considerably more time.

CHEATING. Ten students out of 187, mainly under pressure of the deadline, tried to cheat. Possibly this was a generalization of a cultural pattern in trying to beat any machine. Seven tried to cheat by writing each answer one frame behind the presentation of the item. This practice could be identified by the first frame punches on the response strip. The other three either copied answers before inserting the disk in the machine or worked with the disk out of the machine. Since it was necessary to fake a few wrong answers, all such practices could easily be spotted. Although the disk told the student the correct responses, it gave no help in arriving at the commonest or most plausible wrong answers. (As described elsewhere in this report, a revised model of the machine will make cheating practically impossible.)

Student Questionnaire

When all work on the machine had been completed, the students were asked to fill in a questionnaire and to volunteer additional comments. The most important items gave results as follows:

COMPARING MACHINE INSTRUCTION WITH STUDYING A TEXT. To the item: "When I read *Science and Human Behavior* after working through relevant material on the machine, I felt that the machine:" 1 percent of the respondents checked "contributed nothing," 37 percent checked "was of some help," and 62 percent checked "made the text much easier to understand." To the item: "In comparing work on the machine with studying the text, I felt that, *with the same amount of time and effort*:" 32 percent checked "I learned much more on the machine," 46 percent checked "I learned somewhat more on the machine," 7 percent "there was no difference," 12 percent checked "I learned somewhat more from the text," and 3 percent "I learned much more from the text." Considering the fact that the student population was highly selected and contained many juniors and seniors of considerable college experience and high caliber, it appears to be encouraging that 99 percent felt that the machine helped them understand the text and that 78 percent felt that they learned more from the machine than from the text.

ESTIMATE OF VALUE. To the item: "If machines had not been used this year I believe:" 77 percent of the students checked "I would have got less out of the course," 20 percent checked "It would have made no difference," and 3 percent checked "I would have got more out of the course." The 3 percent here probably represent rugged individualists who were never reconciled to machine instruction. The 20 percent who checked "would have made no difference" are roughly the "neutrals" shown in the following questions.

FUTURE PREFERENCE. To the item: "If I were to take another introductory course in a science or similar field I would:" 67 percent checked "prefer to have machines used for part of the course." However, 16 percent checked "prefer not to have machines used," and 17 percent checked "not care whether machines were used or not." It may seem rather grasping not to accept the 67 percent positive result as sufficient. However, the 33 percent of neutral or negative reactions can be largely attributed to imperfections in material and machines. On a separate item, 5 percent indicated that technical flaws in both had bothered them "greatly," and 68 percent reported that they had been bothered "somewhat." Only 27 percent reported not to have been bothered. Until material and machine can benefit from the very obvious improvements which are now possible, it does not seem to be feasible to try to evaluate any more exactly student preferences on the use of the machines. (The objections to material arose mainly from ambiguous frames. These have been identified by an item check of the response strips and will, of course, be almost completely eliminated from the next version.)

MOTIVATIONAL EFFECTS. The frames on a disk were numbered, and the student therefore knew how far he had to go to complete a set. Each response was reinforced by progress toward completion. These conditions, technically called "reinforcement on a fixed-ratio schedule with counter" (see Ferster and Skinner, *Schedules of Reinforcement*, 1957, p. 89 ff.), may partly explain responses to another question. To the item: "When I was not in a hurry to keep other appointments and no one else was waiting for my machine, I felt:" only 26 percent responded with "relaxed with plenty of time to read and study materials." Fifty-seven percent reported that they felt "somewhat anxious to finish," and 10 percent, that they felt "generally pressed for time, but could not say why." It would appear that the design of the machine has overemphasized motivation. One consequence, discovered in interviews, was that the student was inclined to put down a careless first answer, trusting to a correction after having seen the printed answer to clear up the assignment. This is evidently not desirable. Possibly this pressure was responsible for the fact that 37 percent of the students checked an item indicating that they felt they were "missing many opportunities to reflect on material and consider its implications." As will be seen elsewhere in this report, steps have been taken in the redesign to produce a somewhat slower pace and lessened pressure.

OTHER EFFECTS. Several miscellaneous attitudes were checked in the questionnaire. These had been encountered first in occasional interviews. Since the course emphasizes the implications of a science of behavior in managing human affairs, there is a tendency for some students, particularly those in the humanities and some of the social sciences, to resent the implied manipulation of human behavior. Fourteen percent of those responding to the questionnaire agreed they felt they were "being treated like an experimental organism." This may have been nothing more than a reflection of a point frankly expressed in the lectures: The students had been told that the course was being given in an experimental manner. Only 5 percent felt that the use of machines reflected upon their dignity as human beings. Forty-four percent felt that "the instructor was trying to teach me as much as possible with a given expenditure of my time and effort." A check of grades obtained on an hour test showed that more students receiving an *A* or a *B* felt that the instructor was trying to teach them as much as possible. The same students less frequently complained that they were "missing opportunities to reflect on the material."

VOLUNTARY COMMENTS. In general the voluntary comments were strongly favorable. Many students were enthusiastic, said that the machine work was fun and challenging and that they learned a great deal from it. With others, however, instruction by machine was not successful in relieving them of anxiety. They still regarded working on the machine as escape from a threat of failure. Among critical comments were the following: (a) "it was inconvenient to have to come to a special room to study"; (b) "more disks should be provided for frequent review"; (c) "the students should be permitted to repeat the same disk if they felt the need to do so"; (d) "the material needed to be improved, particularly in eliminating ambiguous frames"; and (e) "the machine encouraged incorrectly reporting a response as right." As to difficulty, some students found the material too easy, and others, too hard. Evidently, however, the level was fairly happily chosen.

ERROR ANALYSIS. In the course of the term several hundred thousand responses were written by the students. Facilities do not permit a complete item analysis, but a sampling has been made. The number of correct first answers and a tabulation of the commonest mistakes or alternative answers are valuable in revising material. Common mistakes also suggest the development of special techniques of composing items to deal with unexpected types of response.

An Application of Recent Developments in Psychology to the Teaching of German

CHARLES B. FERSTER and STANLEY M. SAPON

INTRODUCTION

PRESENT-DAY educators are faced with problems of ever-increasing numbers of students and a curriculum which is continuously expanding in complexity and scope. At the same time, our economy is not supporting these demands with adequate numbers of teachers and sufficient sums of money.

Recent developments in psychology, as presented by B. F. Skinner (1954), suggest some solutions to the need for greater efficiency in the educational process. Skinner's work contains several important principles which could be applied immediately to the field of education. This paper describes the application of these techniques to some of the problems encountered in teaching German composition.[1]

The major principles set forth by Skinner are as follows: By a process known as "reinforcement," new forms of behavior may be created with a degree of subtlety not possible heretofore. The salient feature of this process is making a reward immediately contingent on *small* units of the learner's behavior. Tech-

---

This article originally appeared in the *Harvard Educational Review*, Vol. 28, 1958. This work was carried out while the authors were at Harvard University, and the project received support from the Department of Psychology and the School of Education. The instructional material was designed by C. B. Ferster, who is now at the Indiana University Medical Center. The evaluation was carried out by S. M. Sapon, who is now at The Ohio State University.

[1] "Composition" is used here in the sense of "controlled composition," i.e., the writing of complete sentence material in German whose content is controlled by English stimulus sentences. This is in contrast to the dimensions of written-foreign-language performance, as seen in sentence completion, multiple-choice selection, and free composition. Although free composition reflects the most normal linguistic behavior, it is also the least rigorous, since the student may, by virtue of his free choice, avoid the forms and constructions that he does not command readily. Implied as well is the counterpart to the active skills described above, the ability to read the foreign language.

nically, such a reward is called a reinforcement. By differentially affecting minutely different forms of behavior, it is possible to progress from one form to another in small steps so that there is development from simple to complex forms. An important by-product of the development of complex forms by small steps is that the motivation of the student is not jeopardized by frequent failures. Theoretically, a set of instructional materials could be graded so that the student could progress to complicated behavioral repertories while making almost no errors. Although most current educational theory is based on the principle of reinforcing or rewarding the desired behavior of the student, much classroom behavior is 'aversively" controlled, e.g., by fear of failure. And even when this is not true, the rewards and reinforcements are postponed beyond the time when they would be maximally effective because of the present structure of the classroom situation. Also, the relation between the reward and the student's activity is too vague to differentially affect specific behaviors. For example, telling a student that 50 percent of an exercise is correct does not differentially strengthen the correct behavior.

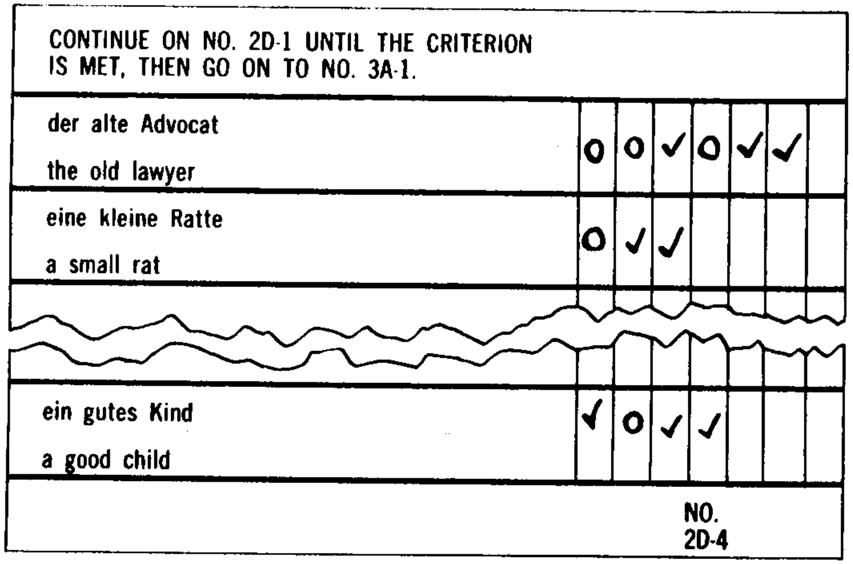
A large amount of training is needed for acquiring such complex behavioral repertories as a new language or an arithmetical operation such as long division. The number of reinforcements required to establish such complicated repertories is extremely large and, in terms of time and classroom organization, beyond the scope of the teacher. These reinforcements are normally supplied by the student himself with the aid of conventional educational materials and the teacher's guidance. The large role played by the student himself undoubtedly accounts for much of the variability in his performance. A set of materials or devices is needed that will provide the intimate and precise interaction between the instructor and learner . . . [that is necessary] to develop complex behavioral repertories. Homework is one attempt to supply this need, but in homework no reinforcing agency can act with respect to small enough units of behavior. The student's behavior is not shaped to create new forms that are more complex than those already existing in his repertory.

The degree of success in the study of a subject such as German composition depends largely upon the study techniques which the student already has. Many differences in study habits among students could be minimized by a detailed specification of how the student is to study the material. Procedures should be built into the instructional material so that each student automatically derives maximum benefit from it. This idea does not necessarily imply a rigid mold to which all students must be fitted, for these procedures can be designed to accommodate differences among students, such as how much time different students (who vary in aptitude and intelligence) will require to learn a particular item.

These principles will be exemplified here in a set of instructional materials in German composition. The method could, of course, have been applied to the other aspects of German language instruction.

DESCRIPTION OF MATERIALS AND METHODS

The instructional material is composed of sheets of paper on which are printed pairs of sentences, equivalent in meaning, in German and English. The sheets are used with a mask which permits exposure of one line of material at a time as well as the scoring column, which is always visible. The student exposes a line of the material presenting an English sentence, and writes on scrap paper the German counterpart. He then exposes the next line of the instructional material, which contains the correct translation of the English sentence. If what the student has written and the printed German conform exactly, the student takes credit for a correct response and goes on to the next item. If there is any discrepancy between what the student has written and the second line of the materials, the student re-covers the second line and repeats the process. Figure I is a sample of the actual instructional material which we will use to illustrate the method more concretely. Each item is composed of two lines, with English on the lower line and the German text on the *other* line. Opposite each item are seven column spaces, which are used to record success or failure. The mask is used to expose one line of the material at a time, leaving the column to the right uncovered. When the student's answer and the text on the next line correspond exactly on the first attempt, the student places a check in the first column to the right. If there is any discrepancy, he places a zero. Additional correction attempts are not scored. The student proceeds in this manner item by item through the unit. At the end of the unit



CONTINUE ON NO. 2D-1 UNTIL THE CRITERION
IS MET, THEN GO ON TO NO. 3A-1.

der alte Advocat the old lawyer	O	O	✓	O	✓	✓	
eine kleine Ratte a small rat	O	✓	✓				
ein gutes Kind a good child	✓	O	✓	✓			

NO.
2D-4

FIG. I. Sample of instructional material

The mask is moved upwards, exposing first a line of English and then a line of German, and so on.

he is instructed to return to the first item in the unit and to continue the process as before. Successes or failures for this second trial run are noted in the second column to the right. The student repeats this process until every English sentence has been exactly translated into German in *two consecutive* attempts. The student can scan the scoring columns at any time and determine which items he can skip because they already have two consecutive check marks, working only on those items which he has not yet mastered. The scoring columns of the sample page in Figure I reveal that on the first item, at the bottom of the page, the criterion was reached after four trials. On the second item, the criterion was reached in three trials, and in the third item, six trials were required. The conformance of the student's written German with the material he uncovers on the second line is the reinforcement that maintains the study behavior. The successful completion of a line of the instructional material is reinforcing because it indicates another step toward mastery of the German language.

APPLICATION OF PRINCIPLES IN THE DESIGN OF THE INSTRUCTIONAL MATERIAL

Psychological principles other than those of immediacy and the specificity of the reinforcement appear in the design of the sentences in the instructional material.

1. The amount of work per reinforcement is kept low. The occurrence of reinforcements during the course of the instructional material depends on the amount of work the student does. Reinforcements programmed on the basis of amount of work done have special effects which have been studied extensively in lower organisms. It has been found there that the organism shows a lessened disposition to return to work as the amount of behavior (work) required to obtain reinforcement is increased. This lessening of motivation continues until at extreme values the behavior will no longer be emitted. The decreased motivation is not caused by physical exhaustion, because under similar circumstances equivalent amounts of work will be done without signs of fatigue. Every time a student translates a line of the instructional material incorrectly, the amount of work he does per reinforcement is increased and his motivation is correspondingly decreased. Any factors which will reduce the number of errors, therefore, will keep the ratio of reinforcement to work high and avoid the lessening of motivation that occurs as a result of too much work per reinforcement.

2. Vocabulary. New vocabulary is introduced in a controlled manner. After each new word is introduced, it is included in the subsequent few units in order to provide practice. No attempt is made to match the natural frequency of occurrence of the various words. Words which occur very infrequently in the natural language are used as frequently as those words which occur much more frequently.[2] This prevents overlearning of

---

[2] This is an attempt to overcome a serious shortcoming that derives from too rigid an adherence to frequency word counts. Where comprehension is vital, the lack of understanding of a single low-frequency item may yield unfortunate results. The criticism here leveled does not concern the use of frequency lists as a guide for syllabus vocabulary (frequency was considered in selecting the vocabulary for the present ma-

high-frequency words and underlearning of low-frequency words and minimizes the amount of work required for a given amount of progress in the language. In the construction of the material, a tally was kept of the use of each word so that all of the vocabulary was used as equally often as possible.

3. Control of overlearning. The student continues to work only on those items which he does not yet know. Items which are mastered to the required criterion are omitted.

4. Conceptual terms. All of the conceptual and more complicated parts of the language are taught without any explicit mention of the grammatical, syntactical, or conceptual principle. This is accomplished by making the student translate sentences which require the application of the principle. When the student encounters a sufficiently large number of instances in which the only *new* factor is the principle to be learned, and in which he attempts a translation in which he is variously right or wrong, he comes to behave in terms of the conceptualization even before it can be verbalized.

5 (a). Graded level of difficulty. A graded level of difficulty was attempted so that the progression from item to item is so slight that the student seldom fails. In this way the student makes constant progress in the mastery of the language. Experiments from other fields in psychology show that the motivation of the learner is largely determined by the over-all frequency and the pattern of successes and failures. Ideally, as noted above, a set of materials could be constructed in which the progression from item to item is so gradual that few failures will occur.

(b). Whenever a new principle, vocabulary item, or usage is being learned, it is the only thing being learned at the time. All the other parts of the material have been mastered previously. The second stage in learning these new materials is to have the student use them in varied contexts. For example, the student cannot be assumed to have a thorough mastery of the concept of the genitive case until he has worked with material that requires him to distinguish between the genitive and dative cases. This practice conforms to the basic process of concept formation, in which a crucial element is an opportunity for the organism to behave inappropriately in respect to the concept. For example, if a hungry pigeon pecks at a green disk because the response is reinforced with food, we cannot be sure that he is attending to the green light until other colors are presented and the bird's pecks go unreinforced.

(c). Sufficient practice and training are given so that older material is thoroughly mastered before new material is introduced. The principle of continuous mastery is particularly important in learning cumulative skills such as a foreign language. A temporary lapse of effort or attendance in an early stage of a course frequently makes it difficult, if not impossible, for a student to resume his studies with the class as a whole at a later point. This common source of difficulty is less likely to occur with the method discussed here because the rate at which the student is exposed to new material depends on his mastery of the prior material.

THE DESIGN AND ADMINISTRATION OF THE INSTRUCTIONAL MATERIAL

The instructional material contained 522 words: 286 nouns, 143 adjectives, 72 verbs, 7 adverbs, and 14 prepositions. In addition to these, structural morphology was developed around the article, the interrogative, the pronoun, and the notion of the morpheme (particularly for compounding in verbal and substantive forms). In terms of syntax, the case system with its structural relations was thoroughly presented, as well as exact sentence word-order.

The content of the material closely approximated that of a first-semester college course in German.[3] However, the net recog-

---

terials), but rather relates to the fact that most textbooks tend to teach high-frequency words to a high level of learning, leaving the low-frequency word few contingencies for reinforcement. In terms of probability of evoking the desired response, differential training tends to yield high probabilities of recognition of high-frequency words, and low probabilities of evoking correct response for low-frequency items. It would seem logical that, once a number of vocabulary items have been judged necessary in a text, effort should be made to make the probabilities of correct response equal for all the items.

[3] A typical text designed and used for first-semester college course work in German, *Beginning German Revised*, Otto P. Schinnerer (Macmillan 1935 and 1956) contains 500 items of active vocabulary, as compared to 522 in the present material, and almost identical coverage of grammatical material, such as gender in nouns, pronouns, and articles, verb tenses in the indicative and subjunctive, declension of nominal and adjectival forms, interrogative, reflexive and relative pronouns, etc.

nition vocabulary was smaller, since supplementary reading materials were not used.

The instructional material was administered to 28 subjects who responded to the following announcement posted in many of the buildings on the Harvard campus:

> Graduate students who need or want to acquire a reading knowledge of German may obtain free instruction in German via a new teaching technique currently being developed.
>
> Preliminary results indicate that individuals of varying ability can, with from 15 to 25 hours of study, achieve a level of accomplishment in reading comparable to a full semester of regular college course work.
>
> Students with no previous training in German who are interested in obtaining instruction can do so at hours convenient for them.

The individuals who responded offered a variety of backgrounds: four undergraduates, five law students, 17 graduate students, one secretary, and one housewife. None had any prior training in German. They were informed of the purposes and nature of the investigation and were given three tests of aptitude for foreign language learning.[4] These tests have shown much promise as prognostic instruments with the adult samples in traditional courses studied by the Harvard Language Aptitude Project. The first step in instruction was a demonstration of the use of the materials, followed by a practice run in which 30 German words from the first unit were presented by the voice of a native speaker of German via tape recorder. This was done in order to forestall as much as possible gross mispronunciations, which might handicap the student's future progress in German. The subjects were then given the first part of the lesson material (an estimated eight hours' work) and were instructed to keep an exact record of the dates and starting and finishing times for each unit of work. Subjects were free to do the work at any time or place they saw fit, and at no time were they provided with supervision or instruction by a teacher. When the initial units were completed and returned, the rest of the material was distributed.

ASSESSMENTS OF RESULTS

The student mastery of German composition after completion of the instructional material was determined by his performance in the following tests:

1. A test of vocabulary on a recognition basis presenting 50 German words sampled from the material which were to be identified by a writing out of the corresponding English word. Items were scored right or wrong.
2. A test of the ability to write German sentences by the translation of English material consisting of 12 sentences totalling 100 words in German. These sentences were randomly selected from the last four (most complex) units of instructional material. Items were scored in terms of accuracy of word order, spelling, capitalization, correct article, and verb and noun endings. A failure in any one of these constituted a scorable error. For example, a word to be scored as correct must be morphologically exact, in the correct position, and spelled correctly.
3. The above also served as a measure of active vocabulary independent of structural mastery. This measurement is derived by scoring the items only in terms of the word root, ignoring syntactical considerations and minor spelling errors.

These tests, designed to measure mastery of a given corpus, were felt to have greater specific validity than currently used standardized tests, inasmuch as the latter necessarily attempt to match current textbook and classroom materials and procedures.

Inasmuch as the material represents a large and adequate sample of the German language, completion of the work sheets is in itself a meaningful measure of accomplishment, since in reaching criteria in each unit, the student has, in fact, written the correct German sentences from the English stimulus material.

Some estimate of the efficiency of this method of instruction

---

[4] These tests, developed on the Harvard Language Aptitude Project, consisted of J. B. Carroll's *Paired Associates Test* (a test of rote verbal memory), Carroll's *Words in Sentences Test* (a measure of a sense of functional grammar), and Sapon's *Spelling Clues* (a measure of an ability to recognize and define English words on the basis of minimal clues).

Carroll

is also given by the amount of time spent by the subjects. The largest individual differences occurred here.

RESULTS

Table 1 ranks the six subjects who finished the lessons in terms of achievement on content tests. The table contains the total number of hours spent on the material, as reported by the subjects. The mean time spent on material was 47.5 hours. The range of scores in recognition vocabulary was 76 to 98 percent, with a mean of 88 percent; sentence translation 70 to 93 percent, with a mean of 81 percent; and active vocabulary, 90 to 100 percent, with a mean of 96 percent. The mean scores on the aptitude tests and the rank order of these scores, compared to the ranking of subjects in terms of achievement are also shown in Table 1. It is interesting to note that these aptitude measures, which have been successfully used to predict achievement in traditional courses,[5] would seem to be much less effective in the present methodological context.

Table 1. **RANKING IN TERMS OF ACHIEVEMENT ON CONTENT TESTS***

STUDENTS	VOCABULARY RECOGNITION	SENTENCE TRANSLATION	ACTIVE VOCABULARY	MEAN ACHIEVEMENT SCORE	HOURS SPENT	RANK ORDER ON APTITUDE SCORES	MEAN APTITUDE SCORE
J.M.	98	93	97	96.0	39	1	96.6
R.G.	98	90	100	96.0	78	2	95.3
M.K.	88	83	95	88.6	38	6	54.3
G.R.	76	80	100	85.0	49	5	75.3
E.C.	90	70	93	84.3	38	4	76.3
D.M.	80	70	90	80.0	43	3	89.6
Mean	88.3	80.9	95.8	88.3	47.5		

*All scores in percentages.

---

[5] Carroll, J. B.; Sapon, S. M.; and Richards, S. E. "Construction and Validation of a Test Battery for Predicting Success in Spoken Foreign Language Courses, Report No. 1." The Harvard Language Aptitude Project, June 1954, mimeographed; and Carroll, J. B., and Sapon, S. M. "Cross Validation of Foreign Language Aptitude Tests with a Second Air Force Trial Course Sample, Report No. 2." The Harvard Language Aptitude Project, February 1955, mimeographed.

DISCUSSION OF RESULTS

With a mean time of 47.5 hours, the six subjects learned an amount of German comparable to that presented in a first-semester course. In a semester course, however, the students spend an average of 48 hours in class with their instructor, in addition to the amount of time recommended and spent for homework.[6] Also, the vocabulary and syntax that were learned were active in contrast to the largely passive vocabulary acquired in a conventional course.

The subjects had no instructor and were given no formal statements of grammatical principle; yet, the material succeeded in teaching inductively such conceptualizations as gender, verb transitivity, morphology and syntax of the German case system, and sentence word-order.

MODIFICATIONS IN THE INSTRUCTIONAL MATERIALS DICTATED BY THE PRESENT RESULTS

1. Level of Difficulty of the Materials

While administering these materials, we discovered that level of difficulty was the biggest determinant of the student's sustained motivation. Whenever the level of difficulty increased too rapidly from unit to unit, so that large numbers of errors were made, the reaction was uniform in nearly all the students, who reported being emotionally upset and who reported a strong tendency to stop even though they may not have completed all of the work. What appeared before the experiment to be very smooth progression proved to be much too difficult for the student. Those students who failed to finish the course stopped at the more difficult lessons. We assume that the uneven difficulty level of the instructional material affected the student's motivation by increasing the number of errors and hence the amount of translation the student had to do for each reinforcement. The disposition to return to the study material probably decreased as the larger number of errors produced a situation in which the student emitted too much study behavior per reinforcement to sustain further study.

---

[6] Estimated on the basis of a 16-week semester with three hours of class meetings per week. The amount of time spent on homework can well be considered an extremely important point. Many schools recommend the principle of two hours of preparation for each hour spent in class, and the size of the homework assignments is often likely to require the full recommended time. If, then, a figure approximating 96 hours of outside preparation is considered in addition to class periods, the time differential becomes even more striking.

2. Amount of Work Per Unit Progress in the Language

One of the problems in this particular technique of instrumenting teaching is that each line of the lessons becomes longer and contains a larger number of words as the student progresses from lesson to lesson. Therefore, the amount of gross labor involved in a given amount of progress in German increased with each lesson. This increase in the amount of work is, of course, not necessarily related to the new principle being taught, although it provides practice in the material already learned. In this type of teaching, however, the amount of practice is specified and designed into the material and should not accrue haphazardly. Future modifications of the materials should include other forms, such as partial sentences in German. Here, the bulk of the sentence would already be translated; only those parts involving the vocabulary, grammar, or usage currently being taught would be omitted.

Two points deserve further discussion here. The first concerns the fact that the estimate of "15 to 25 hours of study" made in the posted announcements proved to be too low. Since this estimate was reached by extrapolating the results of several trial runs made with the first half of the instructional material, the effects of uneven progression in level of difficulty and accumulated overhead on working time were not apparent.

Secondly, it will be noted in Table 1 that the total time spent by subject R. G. was considerably in excess of J. M.'s with whom she was closely ranked in terms of aptitude and achievement. Examination of the work sheets reveals that R. G.'s 78 hours were spread over a longer period of time than that taken by the other subjects, with longer intervals between work periods. There may be an implication that long lapses between work sessions have an adverse effect on economy of time without similarly affecting final achievement. This possibility suggests controlled experimentation with this type of instructional material under different conditions of spacing.

3. Shaping New Behavior in the Student

The design of the materials could be improved if each item made the correct answer in the next item more probable. The level of difficulty from item to item could be decreased by the use of all those principles of behavior which determine the form of a particular word response. For example, a series of items could be designed . . . [so] that a new word never before used is made more likely to occur. The word *Fabrik* in response to the word factory could be made more probable by a preceding item, such as "He saw the red fabric."[7] Of course, later items would include the word *factory* where it was not preceded by the word fabric. As a *tour de force*, a series of materials could probably be constructed in which each item is scientifically designed so that the student will progress from a zero knowledge of German to a complicated repertory of the level of a year of college German without ever having made an error. An achievement of this kind would be made possible through use of processes by which new verbal behavior is created rather than by the traditional processes of recall and verbal memory.

FUTURE COURSE OF DEVELOPMENT

Although the present findings are only a first approximation and a preliminary instrumentation of the psychological principles being applied, the results show the possibility of teaching basic foreign language skills with high efficiency and economy of time for both student and teacher. Materials such as these could be used either for introductory preclass instruction so that all students could be raised to a given level of ability before they met their teacher, or they could be used as materials supplementing classroom instruction, leaving the teacher free for those tasks requiring the subtle abilities of the teacher as a human being.

See comment on this paper by G. I. Brown and H. L. Hodgkinson (1958) in Appendix I.

[7] This is an example of a formal prompt. (See Skinner, 1953. p. 213-14, 245.)

A Report on Instructional Devices In Foreign Language Teaching

DOUGLAS PORTER

I. INTRODUCTION

It is the intention of this report to make very specific recommendations concerning the use of devices in language instruction. However, usefulness of the recommendations rests to some extent upon contingencies not yet established; e.g., the time schedule to be followed, amount of money to be put into development of devices, availability of personnel to program teaching materials for the devices, and exact use of the devices. For this reason, the report will attempt to present several alternative courses of action appropriate to the decisions that eventually will have to be made. Yet, since it is felt that one of the functions of this report is to influence the decisions which will be made in developing a language instruction package, the author will not attempt to hide his own opinions on this matter.

II. THE ROLE OF TEACHING DEVICES

There are two major functions of teaching devices in language learning: instruction and research. Although these two functions are closely related and feed upon one another, the type of device chosen and the way it is used may differ considerably, depending upon whether research or instruction is the major interest. In addition, there are several different roles which devices could play in the instructional process. The role to be played will also have bearing upon the type of device chosen. These matters will be discussed before approaching the more technical part of the report.

The first issue is: How much of the instructional process should be undertaken by devices of one sort or another and how much should be left to the usual student-teacher interchange. There are at least two facets to this question: One involves the technical feasibility of using devices, and the other is a matter of strategy and hoped-for payoff. This strategic element is illustrated by this example: There is currently a poor supply of teachers of Russian, and there will continue to be a poor supply because of the difficulties in training good people and because of the small number of prospective teachers available for training. If it is desired to have as great an impact as possible upon the schools as quickly as possible, the larger the proportion of the teaching process undertaken by devices, the sooner the impact will be felt. In addition, it is likely that unless teaching devices take an important position in the program, the result will be similar to past attempts at influencing the course of education, which have depended upon recruitment appeals and conventional training techniques and have resulted in no great impact. To put the matter in another way: In concentrating a large amount of effort upon the use of devices, the element of risk is greater than it would be if a conventional program were followed; but the potential payoff is also greater.

Although actual, tangible evidence for the effectiveness of teaching devices is meager at this point, available evidence shows that "teaching machines" seem to make a valuable contribution to college learning in a psychology course (Skinner and Holland, 1958). Similar machines used by Porter (1958c, 1959) show a consistent advantage over conventional techniques of teaching spelling, both in terms of the quality of learning and in terms of efficiency (amount of learning per hour of machine study). To this evidence may be added the generally favorable reports of the users of language laboratories (Porter, 1958b).

Against the positive arguments above, several factors must be balanced. There is, of course, the matter of risk, which has already been mentioned. The second problem is whether it is possible to produce or purchase teaching devices which will meet the specifications necessary for undertaking the entire process of language instruction. With possible exception of training in the more sophisticated aspects of pronunciation, provision of suitable teaching devices is *not* a problem, and the choice of one type of device over another rests primarily upon training as opposed to research interests of the project. (A more

---

The original version of this paper (1958a) was prepared by Porter at Harvard University for the International Training and Research Program of the Ford Foundation. The editors have deleted several portions which appeared to be already adequately covered by other papers appearing in this volume and have introduced other minor editorial changes. Certain specific equipment recommendations and a discussion of the comparative operation of phonographs and tape recorders have also been deleted.

Ford Foundation

detailed discussion of the problem of training in the spoken language will appear later in this report.) The final problem is that of providing an adequate staff of competent persons to program and prepare the special types of materials needed for device-mediated instruction. Three types of personnel are required for this work: (a) content specialists, (b) psychologists with a knowledge of instructional materials, and (c) technical personnel and native-language speakers who would actually prepare the materials. Content specialists and technical personnel should be readily available. The list of psychologists who have experience in the programming of teaching materials continues to grow from year to year; however, this personnel category would be most difficult to fill. In addition, it is necessary to provide the time and facilities that would allow more or less continuous contact between the psychological and content personnel over a period of time sufficient for preparation of a syllabus and a complete set of instructional exercises.

If the requirements of suitable teaching devices and staff to prepare teaching materials can be met, as they undoubtedly can, then the greatest payoff should come through as complete a utilization of the devices as possible. However, if the teaching devices are looked upon as an ancillary part of the language package, they are in danger of falling into the psychologically unsound pattern of use which characterizes much "language-laboratory" teaching (Porter, 1958b) and which has probably been responsible for the haphazard and generally unproductive use of the "stimulus-type" of teaching devices in the public schools (Porter, 1957). The recommendation of this report, then, is for as complete a utilization of teaching devices as possible: (a) because of the potential contribution they offer to a more efficient learning process; (b) because the usual present-day methods of attempting to break the manpower "bottleneck" are not adequate to the task (several generations of teachers and students will be required before success is possible); and (c) because there are possibilities of transferring any findings and methodologies to other training areas.

III. TYPES OF TEACHING DEVICES

Since the inception of the United States Patent Office hundreds of patents have been issued for teaching devices of one sort or another. Some have been marketed, but a great majority have never reached the production stage; and few of the devices that have been marketed have had any significant influence upon instructional procedures. There are many reasons for this, but one of the most important is that the devices have been poorly equipped to undertake fully the instructional process. It has been pointed out elsewhere that teaching devices may be classified into three major categories: "stimulus devices," "response devices," and "stimulus-response devices" (Porter, 1957). . . .

In foreign language teaching, use has been made of auditory stimulus devices for the most part. Typically, students are required to listen to records or tapes until they think they understand them. Only occasionally are students required to respond actively to the material presented. Although relatively inexpensive and reliable equipment is available which will allow a student to listen and respond and then check on his responses, it is not in very wide use and is employed in a haphazard manner.

The most recent interest in "teaching machines," occasioned by the work of B. F. Skinner, has led to the design of numerous devices of the stimulus-response type which are capable of being adapted to a wide variety of instructional tasks, including language teaching. [Here the author outlined the characteristics of some devices described elsewhere in this volume by Skinner and others.]

. . . .

Because of the large number of commercially available audio-visual devices and the increasing number of new stimulus-response devices, we are left with the problem of selecting, assembling, or designing a device suitable for foreign language instruction. The practical selection of such a device is a complex matter into which many variables enter, but at this point it will be profitable to specify what seem to be its ideal characteristics. Certain of these characteristics will have to be sacrificed to expediency, but it is advisable to realize what exactly is being sacrificed.

First of all, an ideal language-teaching device should be capable of dealing with two major aspects of language performance, the productive and the receptive. "Productive language performance" is defined as the speaking or writing of a language; "receptive language performance" is the understanding or translating of a language both in its spoken and written forms. This is an important distinction to make in terms of a teaching device. If a device is to be effective in teaching the writing and speaking of a foreign language, it must be able to evaluate the student's responses in that language, written or spoken, as well as being able to present stimuli that are appropriate to the responses required. This means that the device must receive or record both the spoken and written output of the student, evaluate it, and feed this evaluation back to the student so that, if necessary, he can alter the form or precision of the new responses he is learning. As in the learning of any new skill, it is generally necessary for the student to repeat certain of the new responses several times in his attempts to reach perfection. An ideal device would allow this repetition. On the other hand, learning the receptive aspects of a foreign language involves, not the learning of new responses, but the attaching of old responses (native-language responses) to new *stimuli;* for example, responding "dog" to the new stimulus *chien*. Again, the teaching device must be able to evaluate the student's responses, but these responses can all be made in a written form which is much easier to handle than are spoken responses. Secondly, an ideal teaching device should be capable of presenting a predetermined sequence of stimuli to the student, with the option of breaking the sequence at certain points for repetition of particular presentations.

To describe all of this in other terms, an ideal language-teaching device should be capable of presenting stimuli and evaluating responses in both visual and auditory form. It should be capable of presenting a particular sequence of stimuli, with the option to repeat certain items in the sequence left to the discretion of the student. (Of course, such a device should be trouble-free, inexpensive, small, have high fidelity, be easy to prepare materials for, be readily available.) The closest approach to these specifications is currently available through the combination of certain magnetic recording-playback equipment with the newer Skinner-type of stimulus-response teaching device, which is designed to handle graphical (printed and written) presentations and responses. In the section to follow, the many practical and theoretical factors which should be taken into consideration in recommending particular devices will be delineated.

IV. THE SELECTION OF A TEACHING DEVICE

A number of factors which influence the selection of a teaching device will be considered and a specific recommendation made concerning each. In some cases there will be an obvious interaction between different factors that can lead to conflicting recommendations. For instance, how expensive is it to obtain a certain degree of audio fidelity and how important is this degree of fidelity in teaching accurate pronunciation? Insofar as possible, recommendations will be made in respect to the aspect of conflict that seems to be most crucial, but final decisions will have to be determined by contingencies unknown at the time this report is being written. The various factors will be considered in an order which approximates their importance in obtaining the best possible devices.

A. Machine Functions

If teaching devices are to take over a large portion of the language-teaching process, then it is crucial that they be capable of presenting both auditory and written stimuli, as described above. Equally important is the requirement that the devices give a student immediate confirmation of the correctness of his responses. . . . (For further discussion of the problem of providing adequate response confirmation, see Skinner, 1954, and Porter, 1958c.)

1. AUDIO FUNCTIONS. Many different media are available for presenting the audio portion of a teaching-device program to the student. Foremost among these are magnetic recordings and the phonograph. Each has certain advantages and disadvantages which will be weighed.

Fidelity. Magnetic recordings are generally of a higher fidelity than phonograph recordings, primarily because the preparation and reproduction of magnetic recordings do not involve problems of transducer mass, tracking, and impedance mis-match present in acoustical recordings. In addition, unless exposed to stray magnetic fields, a magnetic recording can be expected to have a longer life at original fidelity than a phonograph record subject to stylus wear. However, phonograph records are capable of more than adequate frequency response and fidelity to meet any language-teaching requirements.

No general statement which will be agreed with by language teachers, psychologists, and engineers can be made about the minimum fidelity requirements for language recordings. Language teachers demand very high-fidelity recording reproduction in order to expose subtleties of pronunciation. Aside from the problem of whether these subtleties can or should be taught in a one- or two-year course, it is quite obvious from a number of studies that a frequency-response range of 200 to 4000 cps is sufficient for high intelligibility. Data reported in *Handbook of Experimental Psychology* [Stevens: 1056; 1951] indicates that an articulation score of 95 percent is obtainable with a frequency range of 400 to 4000 cps. The spoken material in this case was nonsense syllables, quite analogous to a foreign language. If a narrower range frequency response is necessary, loss at the bottom end of the audio spectrum is preferable to loss at the top end because 90 percent of the intelligibility is carried by consonants, which lie primarily in the frequency range above 1000 cps. In an informal study (at The Laboratory for Research in Instruction at Harvard) it has been found that . . .[a sound reproducer] with a frequency range from about 300 to 3500 cps will give articulation scores of about 90 percent when the stimulus materials are such that it is difficult to discriminate syllables, such as *spud, stud, jud, slud, shud, thud,* and *fud.* This should be entirely adequate for language teaching. [A discussion of the relative practical merits of various forms of tape and disk recording was given here in the original report.]

. . . .

Handling vocal responses. For recording a student's vocal responses, magnetic equipment is the obvious choice, both technically and economically. However, there is some question whether an entirely separate recording-playback channel is necessary for the student's responses or whether they can be placed on a parallel channel of the audio presentation equipment. The latter alternative is more economical but will not allow flexible juxtaposition of the student's vocalization with the model vocalization, an operation necessary if immediate confirmation of responses is to be provided. We currently know so little about effective means of teaching adequate pronunciation that the most flexible possible means of presentation should be adopted and an entirely separate recording-playback channel used.

There are only two other alternatives to having a student record and compare his vocalizations with the prerecorded model: having a native-language teacher available or having a device capable of recognizing correct speech through mechanical means. The native-language informants are not available, and neither are automatic speech-recognition devices. It is difficult to say which will be available first but automatic speech recognition, contrary to current newspaper publicity, faces many difficult technical problems. "Audrey," a device developed at Bell Laboratories, which automatically recognizes spoken digits is representative of the state of technical development of such devices (David, 1957). She has a vocabulary of 10 or so words which are recognized with about 90 percent accuracy when "tuned" to certain individuals; when "listening" to other individuals, accuracy drops considerably. Obviously, such devices would be of little use in teaching. The other drawbacks to automatic speech recognition, assuming it were technically feasible in the near future, are bulk and expensiveness of the equipment.

Summary. Assuming that the device should be capable of giving spoken-language instruction without assistance from a native-language speaker and without automatic speech-recognition apparatus, the following components are required: (a) a prerecorded audio channel to present a question, an incomplete statement, a speech model, etc., in the foreign language—this function can be served by commercially available magnetic tape-loop equipment provided with appropriate controls; (b) either a parallel recording-playback channel on the same tape or (preferably) a separate, instantaneously resetting, recording-playback device—the preferred device can be adapted from either commercially available equipment or from special telephone equipment; (c) devices equipped with simplified controls for volume and for tape movement.

2. **VISUAL FUNCTIONS.** The presentation of visual language-learning materials and recording responses to these materials is, in some ways, a more complex problem than the handling of audio materials. Mainly, the problem is that commercially available equipment does not possess the necessary degree of flexibility. However, there *is* commercial apparatus, which, with minimum effort, is adaptable, and specially designed teaching machines, which with certain compromises can be used, *may* be available soon.

Visual presentation. Three methods of presentation will be considered, each of which has its advantages and disadvantages. Moving pictures, according to extensive research (Hoban and Van Ormer, 1950), can have dramatic, highly motivating characteristics and are capable of close realism. However, much of the dramatic effect is a result of what is *done* with the medium, rather than the medium itself; and unless dramatic and dynamic motion effects are needed in language teaching, the technical difficulties of film production should be avoided. Even with the technical simplifications brought about by video tape, producing moving pictures would be a complex business, expensive and time-consuming. If one of the aims of language instruction is to enable students to interact spontaneously in the new language, this can probably be effectively accomplished by structuring real social situations (as has been done at Middlebury and other colleges).

The two remaining media for controlled presentation of textual and pictorial materials are projection of still films (slides or strip) and direct viewing of matter printed on an opaque medium. Here, a choice depends upon the convenience of preparing the materials and the availability of devices for presentation. (See Porter 1958c.)

. . . .

Considering the current unavailability of teaching machines which can use easily prepared materials, and assuming that it would not be desirable to develop a new device, it is suggested that a modified slide projector could be used for presentation. This suggested solution has several advantages. The projector can be equipped with a small opal glass screen for individual use, or the screen could be removed for teaching groups by projection. Several currently available slide projectors are equipped with solenoid-actuated slide-shifting mechanisms which allow synchronization with a spoken presentation. Slides are easily prepared by typing or drawing on cellophane which is then ready for immediate use or immediate reproduction by . . . a slide-maker setup. Slides are stored in magazines, which are easy for a student to load into the machine; the order of presentation can be changed readily—an operation impossible with strip film or continuous printed matter. It is relatively simple to synchronize a slide projector with taped presentation by means of one of the commercially available low-frequency-synchronizer units which operate on an impulse from the recorded material. The disadvantages of using a slide projector are that it would have to be modified and a separate unit added on which students could write their responses. It is not known whether or not this would be as expensive and as time consuming as developing a new teaching machine for the visual instruction, but past experience in developing the simplest sorts of teaching devices indicates that a year is needed (under university conditions) for design, for experimental models, and for ironing the faults out of final models.

Handling written responses. As was the case with a student's spoken responses, it is necessary to provide some means by which written responses may be put down for comparison with the correct response. This step, which provides immediate confirmation of the correctness of responses, is crucial. Current teaching machines are set up in such a way that a student writes his response either on the teaching material or on a parallel paper tape. Movement of a lever exposes the correct answer and moves the student's answer under a glass so that the answer cannot be changed. The student then makes his comparison, scores himself, and brings up the next question. Such an arrangement is more economical if responses are written on the parallel tape and do not use up the teaching materials, but this kind of an arrangement requires more complex mechanisms. For research purposes it is easier to analyze student responses if these are directly on the teaching materials, but relatively simple techniques can be devised to obtain the same data from parallel tapes. If the language programs developed are to be tried out in any extensive manner, it would be preferable to use parallel response tapes and to provide special instrumentation for gathering and analyzing student responses.

If one of the current teaching machines is used in the project, a most likely prospect would be the "general purpose" disk machine used by Skinner. This device has the parallel tape response system. Another alternative would be to make a modified version of the Porter device (Gilbert, 1958b) which would have the advantage of easily prepared materials. A final solution lies in the development of a written-response tape which could be synchronized with the slide projector presentation, using the same means of synchronization as between slide and tape.

This solution, however, would require the design and construction or purchase of the response unit.

Summary. Assuming that the device should be capable of giving visual presentations in sequence and thus provide a means by which a student can record his written responses and obtain immediate response confirmation, the following alternatives are available: (a) visual presentation by slide-projector or opaque presentation device—this function could be served by a modification of commercially available slide projectors, by one of the current teaching machines or a modification of one, or by an extensive modification of the Porter teaching machine; (b) recording of students' written responses could be handled by current teaching machines or by a specially devised paper tape mechanism to operate in parallel with a slide projector; (c) if a parallel tape system is used, some means should be provided to enable convenient analysis of student responses to the teaching materials.

3. COMBINED FUNCTIONS. Although the above visual and auditory functions have been discussed separately, they should be combined in a suitable teaching device. It would be feasible, and perhaps more economical, to provide separate devices for the visual and auditory modes of teaching; but this would involve the loss of effectively combining instructional methods. Very little is known about such matters, psychologically or pedagogically; but it is possible that the effective use of combined visual and auditory stimuli would result in more learning than is commonly expected. Separate devices would also require that students synchronize them when necessary, a procedure that could easily lead to errors and possible cheating. Therefore, it is strongly recommended that any teaching devices provided be equipped in one unit for the entire instructional process.

In the following "block diagram" (Figure I) the essential units of such a teaching device are represented in terms of their interconnections and relationships to the student. The physical layout is roughly one that would be suitable for actual practice, but the diagram itself is meant to be more than a *schematic* (not a pictorial) presentation. Lines drawn between the various components represent electrical and/or mechanical connections necessary for synchronization and student control of the device functions. Depending upon what components are finally chosen, the units represented are the following:

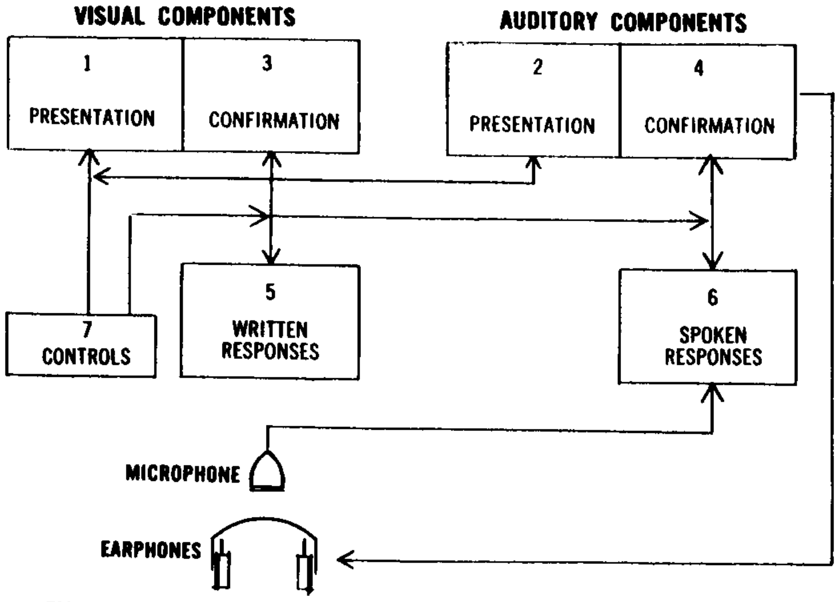


FIG. I. Block diagram showing functional elements of a teaching device

1. Units 1, 3, and 5—Skinner's general purpose machine or modified version of Porter device; or
2. Units 1 and 3—modified automatic slide projector in conjunction with Unit 5—specially designed written-response unit and Unit 7—specially designed controls.
3. Units 2, 4, and 6—dual-channel cartridge-loading endless loop-tape recorder playback unit; or
4. Units 2 and 4—dual-channel cartridge-loading endless loop-tape playback unit in conjunction with Unit 6—special recording playback unit with rapid-reset feature.

Final selection of component units depends upon further considerations to be presented below.

B. The Concepts of Programming and Editing

PROGRAMMING. The term "programming" in the sense it will be used here refers to the process of constructing sequences of instructional material—usually conventional textual material. Instructional materials—whether presented by a teacher or a teaching device—are a crucial aspect of any teaching; thus the ultimate success of any venture such as a self-contained unit of language instruction depends upon the adequacy of these materials. Currently very little is known about techniques of programming. The usual methods involve artful manipulations by good teacher-writers who are dissatisfied with available texts, but, in general, very little can be said about their methods; the manipulations they perform are of a subtle, subjective, and nonverbalized nature.

On the other hand, there are a number of more or less generalizable principles of programming that are derived from reinforcement theory as well as others that have been developed in attempts to write successful instructional programs with teaching machines. The reader should refer to articles by Skinner and Holland (1958) and Gilbert (1958b) for extended descriptions of specific programming techniques and examples of sequences of teaching items. [An outline of these principles was presented at this point in the original report.]

. . . .

It is not suggested that these procedures represent any final word in programming instructional materials or that they are even sufficient for writing a self-instruction language course—for knowledge of programming is meager at this time. However, it is felt that some set of principles such as these, if used in conjunction with the intuitions of good language teachers and the knowledge of language experts, should produce superior teaching materials. It should be emphasized that the use of mechanical teaching devices will make possible the exploitation of instructional techniques that have been seldom (or never) attempted before. Although some of these techniques may oppose current opinions and practices of the language-teaching profession, it would be disheartening to not attempt these techniques, which have sound psychological rationales, simply because of failure to provide personnel with the appropriate psychological training.

EDITING. Closely related to the tentative nature of our knowledge of programming teaching materials is the concept of editing as a research and experimental activity. Traditionally, editing has been a task reserved for publishers who hire individuals skilled in the intricacies of English style and grammar. The activities these editors engage in have little to do with the content or arrangement of content in the texts being edited; content and its organization have been taken care of, presumably, by an author who is an authority in his field and who may sometimes be an outstanding teacher and writer. Programming of the textual material is left entirely up to the knowledge, skills, whims, and biases of the author; and the validity of the materials depends upon the soundness of the author's judgment in regard to how much the textual materials contribute to his teaching; this is a matter usually confounded by other instruction that the students receive in the form of lectures, discussions, and outside readings. However, the use of teaching machines in an entirely, or largely, self-instructional set of lesson materials should allow the authors or publishers of the materials to overcome these editorial difficulties. Student responses to teaching sequences can be analyzed and the materials modified and tried out again as many times as necessary to perfect the lessons. In fact, it would seem mandatory to utilize this research function which is made possible by teaching devices; not to do so would be to pass up one of the greatest long-term advantages of such teaching techniques, the possibility of gradual cumulative revision and perfection of lessons based upon factual findings rather than educated guesses. Such research techniques have paid off handsomely in other areas of technological advance and can be expected to do the same in education.

If it is anticipated that such research-editing activities will be undertaken . . . , then it is mandatory that the teaching devices used, or at least some of them, be equipped to facilitate the research. Such facilitation requires the following features:

1. *Teaching devices which will accept easily prepared and modified sequences of visual and auditory material.* Magnetic recordings are eminently suitable for the auditory presentations, and a modification of the Porter device would provide the most practical visual materials because this device uses standard paper sizes and duplication processes.
2. *Techniques for the easy analysis of students' spoken and written responses.* The analysis of spoken responses

in the form of recordings is an arduous and time-consuming task. One feasible way of speeding up such work is to provide some sort of recorded signal on the tape at the point where error responses lie (such signals could be automatically placed on the record every time a student indicates that he has made an incorrect response). Then, for analysis, the tape could be sent through a playback device at high speed until one of the signals shifts the machine to an appropriate playback speed. This process would be possible through a modification of a standard tape machine. Analysis of written responses presents a similar problem, but it is not as serious a problem because of the speed of human visual scanning in comparison with listening. Basically, there are two techniques that can be utilized: (a) The students would write responses directly on the teaching materials and would score themselves "correct" or "incorrect" by an appropriate mark. Analysis of these resources would proceed by visual inspection; and for 100 or so students, this would not be too cumbersome a scheme. (b) A more automated procedure would require that the responses be placed on a separate parallel paper tape. As a student scores himself, a hole would be automatically punched in the tape next to each wrong response. Then the tape would be sent through a special sensing device which would stop at each punchmark to enable examination of the wrong response. This second scheme would, of course, be more expensive and probably is unnecessary unless there are a very large number of student responses to be analyzed (say approximately several hundred students).

3. *Teaching devices that are flexible enough to allow investigation of all possible combinations of visual and auditory presentation, response, and confirmation within the limits of tape recordings and still pictures.* To provide insufficient research flexibility would be to invite perpetuation of time-honored techniques of teaching which there is some hope of improving.

C. The Possibility of Using Complex Electronic Equipment

There are two ways in which complex electronic equipment could materially aid the language-teaching process: (a) through the provision of devices which would adjust themselves to individual differences in students and (b) by means of special techniques to aid in the teaching of foreign speech. Both would be expensive and cumbersome, and both would require long periods of time for their development if they are to be made operationally satisfactory. A safe rule-of-thumb estimate is that cost, operational difficulties, and development time increase geometrically with complexity of equipment. For this reason it is recommended that unless the language project is willing to devote a number of years to development of equipment, teaching devices be kept as simple as is consistent with obtaining adequate presentation and confirmation of the lesson materials.

If long-term development of special devices is anticipated, there are several techniques which should be considered. In the matter of adjusting to individual differences in students, there are several procedures which can be followed. Teaching materials can be "skip programmed." If the entire lesson involves 2000 items for the student to respond to, the initial stages of the program can require the student to respond to only every fourth item. When few student errors are made, an individual can proceed throughout the entire lesson in the same fashion; but whenever a significant number of errors are made, the device automatically shifts to presenting every second item. In this manner, the better students could go through the material more rapidly than the students with poorer background or ability. Another self-adjusting procedure is one in which the students do all items in a lesson, but the teaching device backtracks and goes over the same material again if errors become frequent. A preferable technique would be to have the machine backtrack to a new and different, perhaps longer and easier, sequence of items covering the same material. All of these techniques considerably complicate a teaching device and probably can be handled alternatively by careful programming of one set of materials or by close administrative supervision of use of the teaching devices.

Giving adequate speech instruction and practice in a foreign language is a more complicated problem, especially the training of "native" pronunciation. The student is usually encumbered by his own native language, which has sounds subtly different from those of the new language and is entirely without other sounds prominent in a new language. The problem is usually complicated further by orthographic differences and irregularities. Often it is difficult for a student even to discriminate important differences in foreign speech because such phonetic distinctions are not made in his native language. In a practical teaching situation such problems are handled by native-speaking instructors who have to work on an almost entirely individual basis with students in order to catch and remedy the student's speaking flaws as they occur. The use of teaching devices may make it economically possible for each student to have an individual "tutor" to handle this aspect of language learning. However, since such devices will depend upon a student's recognizing how good his own pronunciation is when compared with a standard, the procedure is open to errors caused by student misperception of the foreign speech. One obvious solution to this problem is to provide devices capable of recognizing correct speech and informing the student whether his pronunciation is up to par. As was pointed out earlier in this paper, such speech-recognition devices are not available at this time; and although they may become available within the next 10 years, cost, complication, and inconvenience would preclude all but fixed and central location use in teaching.

Other possible instrumentation in spoken-language instruction involves special forms of visual and auditory display of the details of pronunciation. For instance, it is conceivable that a cathode-ray-tube display of the time, frequency, and intensity characteristics of a student's speech could be compared by the student with a simultaneously available visual model of correct pronunciation. Again, such instrumentation would be complex, especially if the device is to make possible an almost instantaneous comparison of the dynamic features of speech, as it must in order to be of much use in reinforcing correct vocalizations. It is also questionable whether students would be capable of reading the complex visual display—language experts have difficulty in reading the static displays provided by a "sound spectrograph" in an already known language. Perhaps a more fruitful type of instrumentation could be provided by presenting and having students attempt to copy speech models modified in temporal and frequency characteristics. For instance, the correct pronunciation of certain sounds or words might be simplified by having the auditory model presented at a slower pace, allowing more time to listen to inner structure of the vocalization. Or filters could be inserted into the speech circuits in order to remove certain components of the total vocalization which tend to mask smaller but important details of pronunciation. Little is known about the effect of these auditory manipulations of language-learning materials so that a good deal of experimentation involving relatively complex equipment would be necessary in order to obtain workable teaching procedures. This type of manipulation of learning materials, however, lies close to the concept of "shaping" behavior that was referred to earlier. One flaw of much spoken-language instruction is that the teacher expects his students to attempt the final, complete, and correct pronunciation from the very beginning, despite interference from former language habits and the complexity of the muscular movements involved in pronouncing new sound combinations. This procedure is almost as psychologically unsound as that requiring a student to translate a page of foreign text *de novo.* What should be attempted, instead, is to lead the student through a series of different pronunciations that approximate the final objective. This process should take into account the particular ways in which the student's native language interferes with the new language, and the sequence of steps should attempt to forestall the usual mispronunciations of students. The above-mentioned techniques for manipulating auditory materials could form an integral part of this type of teaching process; and it is felt that this approach, rather than attempts to use automatic speech recognition, will produce greater payoff.

D. Summary and Recommendations

The recommendations to be made fall into three major categories: (a) the role of teaching devices in the development and the use of a self-instructional language package, (b) the problem of programming lesson materials, and (c) the selection of appropriate devices.

THE ROLE OF TEACHING DEVICES. It has been suggested that the role of teaching devices be a major one because of the potential payoff through application of technology to the educational process. A successful technology of teaching devices could result in economic and manpower advantages which usually have to await the rise of a new generation of teachers trained in the desired discipline—in this case, the Russian language. It is also felt that less than full utilization of devices could lead to dependence upon traditional instructional techniques that are often based more upon personal whim and prejudice than upon psychological knowledge. Despite the great potential contribution of teaching devices to language instruction, this warning is necessary: Teaching devices are no "pie in the sky" to cure the ills of traditional education. We are only in the initial stages of developing a technology of instruction, and much experimentation remains to be done.

PROGRAMMING LESSON MATERIALS. With teaching devices, as in any type of teaching, the selection, organization, and presentation of educational content is a major problem. This is an area in which little systematic research has been done, primarily because teachers responsible for academic content have had neither the motivation nor the skills necessary for the task and psychologists have been more interested in exploring basic behavioral processes than in applying their knowledge to practical problems of teaching. The teaching device provides a research instrument *par excellence* which makes feasible the exploration of innovations in teaching and provides a focal point for collaborative work between academic-content specialists, teachers, and psychologists. This collaboration is long overdue, and the fruits of such an interdisciplinary endeavor could be useful to education as a whole.

THE SELECTION OF APPROPRIATE DEVICES. The selection of teaching devices for the language project rests upon a complex of interacting factors, and the importance of particular factors depends upon exactly what purposes the project turns out to involve. Most important is the provision of appropriate psychological functions: ability to present auditory and visual materials to the student, provision of a means of responding, and confirmation of the student's responses. Next in importance is sufficient flexibility to allow all possible combinations of presentation and confirmation of lesson materials. Directly related, insofar as editing as a research function is important, is the necessity of providing for easily and inexpensively prepared materials to be

used in the teaching devices. Finally, there are considerations of simplicity, reliability, economy, and time. . . . [The concluding portion of the original report presented a brief discussion of some specific devices available at the time of writing.]

Teaching Machines

DOUGLAS PORTER

MEETING THE DEMAND FOR EDUCATION

. . . .

THE WORLD-WIDE demand for education is rapidly outstripping available educational resources. More schools are being built and more teachers are being trained, but this conservative attack is not enough to stem the tide. New ways must be found to improve the efficiency of the educational process itself. While many areas of human endeavor have reaped the benefits of modern technology, education has remained relatively unchanged for centuries. This need not be the case. Technical skills presently available in our society make it possible to implement certain psychological principles which have a profound effect upon learning. Teaching machines represent one attempt to improve the efficiency of education by applying these principles to the teaching process.

Teaching machines are essentially devices for increasing a teacher's control over the learning activities in a classroom. In this sense they may be classified with books, blackboards, motion pictures, and all other teaching devices. However, the teaching machines described here are fundamentally different from traditional devices, in that they are able to evaluate a student's responses to educational content. At one time all teaching was probably done in a face-to-face situation with its constant student-teacher interchange, which necessarily limits classes to a small number of students per teacher. With the invention and commercial development of printing, the influence of a single teacher could be more widely extended through the use of books. Succeeding technological advances have resulted in other more potent methods of disseminating knowledge, but each technological advance has generally led to decreased teacher contact with students and thus left the actual learning process followed by each student more or less to chance. Insofar as technological advances have had this effect, the commonly heard criticism that audio-visual devices are peripheral to education is probably valid. By both presenting educational content and mediating the learning process, however, teaching machines preserve the technological advantages of printing and other devices which allow economical dissemination of knowledge and at the same time reinstate some of the advantages of the face-to-face teaching method.

UNDERLYING PSYCHOLOGICAL PRINCIPLES

The pedagogical advantages of teaching machines stem from effective application of the principles of reinforcement, which have been the subject of much research in recent years. When an organism behaves, it acts upon the environment. This changes the environment in some way. In turn, changes brought about in the environment feed back to the organism and, as a consequence, affect its future behavior. When environmental feedback strengthens the behavior which brought it about, it is said that the organism has been "reinforced" or rewarded. Many events which serve as reinforcers are related to biological processes. Thus food is reinforcing to a hungry animal. Because of their long and complex learning histories, human beings are reinforced by additional types of environmental events which may be only remotely related to biological processes. Thus praise, approval, success, novelty, and many other events are reinforcing, depending upon the individual's particular past experience.

The relationship between behavior and a reinforcement, by virtue of which the behavior is strengthened, is called a "contingency of reinforcement." In order for contingencies of reinforcement to be effective, however, certain conditions must be met: (a) Reinforcement must take place *immediately* after a response has been made. A time interval of more than a few seconds between response and reinforcement greatly reduces the effectiveness of the reinforcement. (b) The reinforcement must be made *precisely contingent* upon performance of the behavior that is being taught. If this is not the case, different and perhaps

---

This article is reprinted, with minor editorial changes and the deletion of a prefatory paragraph, from the *Harvard Graduate School of Education Association Bulletin*, Vol. 3, March 1958. Use of machines to teach spelling, described by the author, was sponsored by the U. S. Office of Education and the School-University Program for Research and Development. One of the figures presented in the original article has been deleted in the present volume.

three factors—how good the textbook is, the influence of other teaching techniques which may have been used, and differing study habits of the students—they cannot reveal *why* certain sections of the test were failed but only tell which students are doing satisfactory work in general. The teaching machine programmer, however, *can* discover good and poor portions of his text by examining student responses to particular items, and he can rewrite those items on the basis of such information. This imposes demands upon the programmer that an author does not face. He is *forced* to think in terms of the responses students will make to specific items; he is forced to analyze carefully the exact behavior he is trying to teach, how it relates to previous learning history of the student and how it will aid in future performance; if certain facts or relationships are to be taught, then he must provide the student with items designed to elicit responses appropriate to these learnings. The end result is that in writing materials for teaching machines the programmer is able to shape up his teaching program to a high level of perfection, a process that the authors and publishers of textbooks are not currently able to undertake.

Self Directed Learner

In preparing teaching materials for current research at Harvard several programming techniques are being utilized: (a) The content to be taught is broken down into a large number of small steps. (b) Successive steps in the lesson are "prompted" by preceding items or by hints provided explicitly for that purpose. (c) Prompts and hints are gradually "vanished" or removed so that by the time a student has completed a lesson he is "on his own." Through these techniques almost all wrong responses are eliminated (in a perfect program a student would make no errors at all), which makes it unnecessary to unlearn incorrect responses at a later time—a notoriously inefficient process. The above principles of programming are what distinguish teaching machine instruction from "rote learning" or "drill." Rote materials presented for memorization are not accompanied by prompts, so the student has to guess at the correct answer. The result is many wrong responses, which may be frustrating and certainly will have to be unlearned. As the drill material is gone over time after time, a student gradually becomes able to provide his own prompts for the correct response. At this point he has had to make many more repetitions of the lesson content than a student using a teaching machine program. In addition, rote learning procedures present unorganized and perhaps unrelated items of knowledge to be memorized, while a teaching machine program is designed to take advantage of inherent organization in the subject matter being taught and to make use of this organization in providing prompts useful for shaping up the student's learning.

It may seem that teaching machines are nothing more than testing devices. This is not so. Although teaching machines could be used for testing purposes, both the machines and their associated educational content are designed especially for teaching. While a teaching machine program contains a carefully arranged series of problems with prompts designed to lead a student gradually to a perfect performance, a test provides only minimal prompts in the questions asked. This use of minimal prompts in a test resembles the final stages of a teaching machine program from which the prompts have been "vanished," after a great deal of learning has already taken place. While the immediate confirmation of responses provided by a teaching machine is crucial for effective learning, it has no function and is not included in a test situation so long as the test's purpose is to evaluate, not teach.

CURRENT RESEARCH

In addition to techniques for the programming of teaching materials several other problems are being investigated: (a) What will the long-term motivational effect of teaching machines be? When first introduced to a class their "novelty value" is high and students like to use them. It is possible that after continued use this effect will wear off. Evidence available to date (after five months of use) gives no indication of lessening student motivation. (b) What kinds of subject matter can a teaching machine effectively present; e.g., factual information, problem solving, critical thinking? Intuitively, teaching machines *should* be capable of handling any type of verbal instruction now undertaken in the schools. (c) What influence on other school instruction will the use of the teaching machines have? One could hypothesize that students will become "dependent" upon teaching machine instruction, making them unfit for other types of teaching. Evidence so far indicates that just the opposite effect is taking place. Young children being taught spelling by machine tend to

X

get rid of poor study habits such as not reading written instructions, looking at only the first two letters of a spelling word, and not looking at what they have written to see if it is correct. There is some indication that these gains in study skill transfer to other areas of instruction, but more evidence is needed.

THE TASK AHEAD

It has been pointed out that new approaches are needed to the problem of providing effective mass education. One solution is to increase the efficiency of the educational process. This can be achieved by effective application of psychological principles of reinforcement to the teaching process. Teaching machines are designed to implement these principles in the classroom more effectively than a teacher is able to and at the same time free the teacher of repetitious tasks which take little advantage of his teaching skills.

The design of teaching machines themselves is a relatively simple task which is fairly well along. Programs of educational content must be prepared and tried out in the machines. This is a difficult and long-term task which has barely been started, one which will actually never be finished. In the light of these facts, teaching machines should not be looked upon as a panacea for all the ills of education, but as devices which hold some promise of improving the instructional process and should be judged, in the long run, by evidence of their success, not by argument.

Please Note!

Teaching Machines: An Application of Principles from the Laboratory

JAMES G. HOLLAND

MUCH HAS been said of teaching machines recently—but the emphasis has tended to be on the gadgets rather than on the much more significant development of a new technology of education initiated by B. F. Skinner (1954, 1958c). The technology does use a device called a teaching machine which presents a finely graded series of problems and provides immediate "reward" or reinforcement for the student's correct answers. But emphasis on machines has tended to obscure the more important facets of the new technology based on application of principles from the laboratory. The machines of today are not necessarily better than those of yesterday. Indeed, adequate machines could have been built hundreds of years ago. The movement today is not simply the mechanization of teaching but, instead the development of a new technology—a behavioral engineering of teaching procedures.

The history of unsuccessful teaching machines illustrates the relatively greater importance of the technique as opposed to the gadgets. The first teaching machine was patented 93 years ago. There have since been a series of patents and a promising burst of activity by Sidney Pressey and his associates (1926, 1927) in the 1920's. None of these early efforts really caught hold. But during this period in which the idea of mechanized teaching has been latent, the science of behavior has developed principles which permit extremely precise control of behavior. This new technology is not only the so-called automation of teaching but is an attempt to obtain the kind of behavioral control shown possible in the laboratory.

The original version of this paper was prepared by Dr. Holland for presentation at the annual Invitational Conference on Testing Problems sponsored by the Educational Testing Service in October 1959 and for publication (1960) in the proceedings of that conference. Two figures and several brief portions of the text have been deleted here because the editors felt the points made were adequately represented by other papers in the volume. The work discussed here by Dr. Holland was supported by grants from the Carnegie Corporation and the Ford Foundation.

as he can, not held back by less able members of the class. Less able students, on the other hand, are not punished by their inability to keep up. Individual machines also allow easy use of special lesson materials for students with particular needs. (c) Two forms of reinforcement are explicitly provided by the teaching machines: Manipulation of the machines themselves is probably of reinforcing value irrespective of the educational content presented; and mastery of the educational content is of still greater reinforcing value, especially since the student is constantly made aware of his progress through immediate confirmation of the correctness of his responses. (d) Since the machines automatically score each response, the teacher is freed from having to correct many papers and may concentrate on arranging discussion, [cf. Figure II], guiding social learning, ministering to individuals with special difficulties, and other activities which cannot be undertaken by machines. (e) A student is not limited in his study by the amount of time a teacher has to guide his learning activities. This has valuable consequences: A large number of responses can be made to a given body of content, and the content can be enriched with topics which there would otherwise be no time to cover.

PROVIDING EDUCATIONAL CONTENT

The mechanics of immediate and precise reinforcement are not difficult to arrange. Without adequate programs of educational content, however, the best devices are of little avail. It is in this area of educational "programming"—the precise selection and arrangement of educational content—that basic knowledge is most lacking and research most needed. It is primarily in this area that psychologists can make a valuable contribution by discovering and applying scientific principles to educational procedures which have heretofore depended upon art alone.

Programming materials for teaching machines is analogous to writing a textbook in many ways, but different in certain important respects. The author, as well as the teaching machine programmer, selects, organizes, and subdivides content into some sort of consistent scheme; he may spread some topics out and skim over others; he may provide examples and other sorts of illustrative material to explain a point. Throughout all of this he strives for clarity of expression in the hope that readers

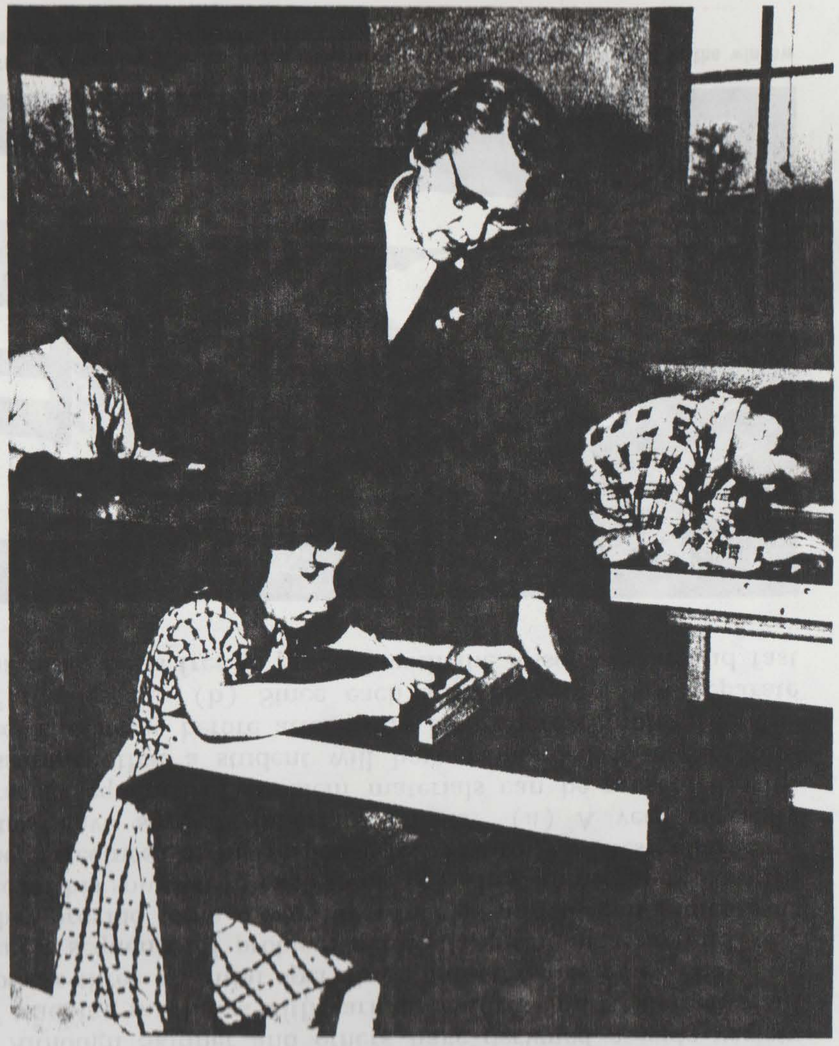


FIG. II. Teacher supervises use of teaching machine as students work at their own rate.

will understand and learn. However, a textbook author never gains very accurate insight into how well his exposition accomplishes its aim. His only source of this information is from tests of the learning progress of students who use the book in a course of study. Since these test scores are a result of at least unwanted behavior will be reinforced. Under these circumstances the wrong sort of performance, or perhaps no consistent performance of any sort, will be taught. (c) A sufficient *number* of reinforcements must be given. Even such an apparently simple task as learning to spell a word involves many separate responses on the part of a student, each of which must be reinforced, at least in the early stages of learning. Failure to provide these reinforcements may result in what is usually called "lack of interest" on the part of the student, or his interest will be shifted to other classroom activities for which sufficient reinforcement *is* available (such as talking, etc.).

By arranging appropriate contingencies of reinforcement, along with other conditions, psychologists have been able to teach or "shape up" complex performances. Similar results have been obtained with a wide variety of organisms, including human beings. Reinforcement as used by the psychologist in shaping behavior is much the same process that a good coach or tutor uses when he points out and rewards desirable portions of a student's performance. By reinforcing a series of successive approximations to the final, complete act, the tutor gradually leads his student to a perfect performance, a process which demands close attention and a delicate ordering of events.

CLASSROOM APPLICATION

Since schools are in the business of shaping behavior—teaching—the above principles of reinforcement apply to teaching activities carried out in the classroom. Skinner has pointed out that in the classroom, however, conditions of reinforcement are largely accidental and not precise simply because a teacher is physically incapable of ministering to 30 or more students all at one time (Skinner, 1954). Providing students with the immediate and precise reinforcement necessary for efficient learning would require a private tutor for each child. For example, an elementary-school child learning to spell 10 words probably has to be reinforced several times for responses to four or five elements of each word. This means that each individual in a class of 30 has to be reinforced roughly 150 times in a 20-minute period. To do an efficient job of instruction, the teacher would have to be capable of reinforcing his class a total of 4500 times, or about 225 times per minute, quite obviously an impossibility.

Schools generally recognize the superiority of individual instruction over classroom teaching by providing a tutor for children with special disabilities. *All* students, however, could profit from the same treatment if it were economically feasible and manpower were available. Teaching machines are designed to overcome these difficulties by providing each student with what amounts to a personal tutor, capable of giving large amounts of immediate and precise reinforcement.

Although Skinner and others have designed a wide variety of teaching machines with various features and refinements, all possess three essential operating characteristics: (a) They present a sequence of problem materials one item at a time. (b) They provide some means by which a student can indicate or record his solution to each item. (c) They immediately confirm the correctness of his response [cf. Figure I]. These characteristics have several important effects: (a) A very carefully graded sequence of problem materials can be presented with assurance that a student will have responded to earlier prerequisite items before attempting later more complex sections of the lesson. (b) Since each student operates a separate machine, he is free to proceed with the lesson as far and fast

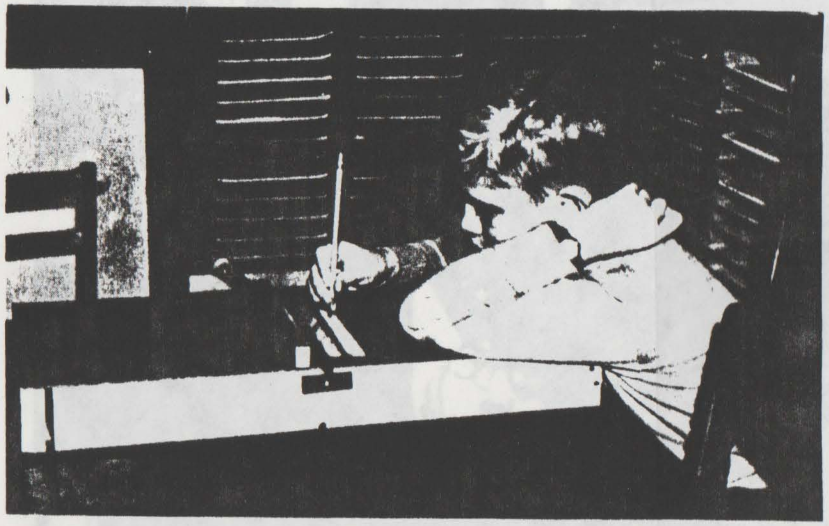


FIG. I. A machine in which lesson materials inserted by the student appear in the window where student writes his answer before exposing the correct answer.

three factors—how good the textbook is, the influence of other teaching techniques which may have been used, and differing study habits of the students—they cannot reveal *why* certain sections of the test were failed but only tell which students are doing satisfactory work in general. The teaching machine programmer, however, *can* discover good and poor portions of his text by examining student responses to particular items, and he can rewrite those items on the basis of such information. This imposes demands upon the programmer that an author does not face. He is *forced* to think in terms of the responses students will make to specific items; he is forced to analyze carefully the exact behavior he is trying to teach, how it relates to previous learning history of the student and how it will aid in future performance; if certain facts or relationships are to be taught, then he must provide the student with items designed to elicit responses appropriate to these learnings. The end result is that in writing materials for teaching machines the programmer is able to shape up his teaching program to a high level of perfection, a process that the authors and publishers of textbooks are not currently able to undertake.

Self Directed Learner

In preparing teaching materials for current research at Harvard several programming techniques are being utilized: (a) The content to be taught is broken down into a large number of small steps. (b) Successive steps in the lesson are "prompted" by preceding items or by hints provided explicitly for that purpose. (c) Prompts and hints are gradually "vanished" or removed so that by the time a student has completed a lesson he is "on his own." Through these techniques almost all wrong responses are eliminated (in a perfect program a student would make no errors at all), which makes it unnecessary to unlearn incorrect responses at a later time—a notoriously inefficient process. The above principles of programming are what distinguish teaching machine instruction from "rote learning" or "drill." Rote materials presented for memorization are not accompanied by prompts, so the student has to guess at the correct answer. The result is many wrong responses, which may be frustrating and certainly will have to be unlearned. As the drill material is gone over time after time, a student gradually becomes able to provide his own prompts for the correct response. At this point he has had to make many more repetitions of the lesson content than a student using a teaching machine program. In addition, rote learning procedures present unorganized and perhaps unrelated items of knowledge to be memorized, while a teaching machine program is designed to take advantage of inherent organization in the subject matter being taught and to make use of this organization in providing prompts useful for shaping up the student's learning.

It may seem that teaching machines are nothing more than testing devices. This is not so. Although teaching machines could be used for testing purposes, both the machines and their associated educational content are designed especially for teaching. While a teaching machine program contains a carefully arranged series of problems with prompts designed to lead a student gradually to a perfect performance, a test provides only minimal prompts in the questions asked. This use of minimal prompts in a test resembles the final stages of a teaching machine program from which the prompts have been "vanished," after a great deal of learning has already taken place. While the immediate confirmation of responses provided by a teaching machine is crucial for effective learning, it has no function and is not included in a test situation so long as the test's purpose is to evaluate, not teach.

CURRENT RESEARCH

In addition to techniques for the programming of teaching materials several other problems are being investigated: (a) What will the long-term motivational effect of teaching machines be? When first introduced to a class their "novelty value" is high and students like to use them. It is possible that after continued use this effect will wear off. Evidence available to date (after five months of use) gives no indication of lessening student motivation. (b) What kinds of subject matter can a teaching machine effectively present; e.g., factual information, problem solving, critical thinking? Intuitively, teaching machines *should* be capable of handling any type of verbal instruction now undertaken in the schools. (c) What influence on other school instruction will the use of the teaching machines have? One could hypothesize that students will become "dependent" upon teaching machine instruction, making them unfit for other types of teaching. Evidence so far indicates that just the opposite effect is taking place. Young children being taught spelling by machine tend to

X

get rid of poor study habits such as not reading written instructions, looking at only the first two letters of a spelling word, and not looking at what they have written to see if it is correct. There is some indication that these gains in study skill transfer to other areas of instruction, but more evidence is needed.

THE TASK AHEAD

It has been pointed out that new approaches are needed to the problem of providing effective mass education. One solution is to increase the efficiency of the educational process. This can be achieved by effective application of psychological principles of reinforcement to the teaching process. Teaching machines are designed to implement these principles in the classroom more effectively than a teacher is able to and at the same time free the teacher of repetitious tasks which take little advantage of his teaching skills.

The design of teaching machines themselves is a relatively simple task which is fairly well along. Programs of educational content must be prepared and tried out in the machines. This is a difficult and long-term task which has barely been started, one which will actually never be finished. In the light of these facts, teaching machines should not be looked upon as a panacea for all the ills of education, but as devices which hold some promise of improving the instructional process and should be judged, in the long run, by evidence of their success, not by argument.

Please Note!

Teaching Machines: An Application of Principles from the Laboratory

JAMES G. HOLLAND

MUCH HAS been said of teaching machines recently—but the emphasis has tended to be on the gadgets rather than on the much more significant development of a new technology of education initiated by B. F. Skinner (1954, 1958c). The technology does use a device called a teaching machine which presents a finely graded series of problems and provides immediate "reward" or reinforcement for the student's correct answers. But emphasis on machines has tended to obscure the more important facets of the new technology based on application of principles from the laboratory. The machines of today are not necessarily better than those of yesterday. Indeed, adequate machines could have been built hundreds of years ago. The movement today is not simply the mechanization of teaching but, instead the development of a new technology—a behavioral engineering of teaching procedures.

The history of unsuccessful teaching machines illustrates the relatively greater importance of the technique as opposed to the gadgets. The first teaching machine was patented 93 years ago. There have since been a series of patents and a promising burst of activity by Sidney Pressey and his associates (1926, 1927) in the 1920's. None of these early efforts really caught hold. But during this period in which the idea of mechanized teaching has been latent, the science of behavior has developed principles which permit extremely precise control of behavior. This new technology is not only the so-called automation of teaching but is an attempt to obtain the kind of behavioral control shown possible in the laboratory.

The original version of this paper was prepared by Dr. Holland for presentation at the annual Invitational Conference on Testing Problems sponsored by the Educational Testing Service in October 1959 and for publication (1960) in the proceedings of that conference. Two figures and several brief portions of the text have been deleted here because the editors felt the points made were adequately represented by other papers in the volume. The work discussed here by Dr. Holland was supported by grants from the Carnegie Corporation and the Ford Foundation.

We have, of course, seen other practical applications of scientific psychology. We are all familiar with the development of a technology of testing which permits placing an individual in situations suited to his abilities. We are also familiar with another technology called "human engineering," which fits machines and jobs to the capacities of man. One places a man in a job that suits him; the other alters the job to suit the man; *neither* attempts to alter or control man's behavior.

For years in the laboratory we *have* controlled the behavior of experimental subjects—both animal and human—by a widening array of principles and techniques. The new technology of education is the application of behavioral laws in modifying or controlling behavior. Such a technology became possible with the realization that we are actually referring to a verbal repertoire (Skinner, 1957e) controlled by the same laws as other behavior. The old, defunct explanatory concepts of knowledge, meaning, mind, or symbolic processes have never offered the possibility of manipulation or control; but behavior, verbal or otherwise, can be controlled with ease and precision.

While machines are not the essential or defining aspect of this technology, they do play an important role in providing some of this fine control the technology requires. We will now examine several machines and notice the advantages they offer.

[The first machine discussed here was the write-in machine described by Skinner, 1958c.]

. . . . . .

Although other techniques such as programmed workbooks "self-correcting" homework,[1] and flashcards are sometimes used in this new behavioral technology, they offer less control. Teaching machines eliminate undesirable forms of responses which would also be successful in obtaining the right answer. For example, the teaching machine insures that the student answer before peeking at the indicated answer. There is a strong temptation to glance ahead with only a poorly formulated, unwritten answer when flashcards are used.

This write-in machine is a prototype of the most common machine. There is another machine used for teaching young children material which consistently has a single possible answer.

---

[1] Diederich, P. B., "Self-Correcting Homework in English." A paper presented at the Invitational Conference on Testing Problems of the Educational Testing Service, October 1959.

In the machine the constructed answer is automatically compared with the true answer. The child is presented a problem, perhaps a statement such as $2 + 2 = x$, and he must provide the "4." By moving a slider appropriately he can insert the 4 into the answer space. He then turns the crank, and the next item appears immediately, and, therefore, immediate reinforcement is provided.

Both of the machines we have seen thus far require the student to compose the answer. Figure I shows a machine for a less mature organism who cannot yet compose an answer. This machine can be used for teaching preschool children (Hively, 1960). There is a large top window and three small windows; in the large window there is some sort of problem, and in the three smaller windows, there are three alternative choices. For example, in the machine, as seen in the picture, the subject chooses one of the three alternatives which has the same form

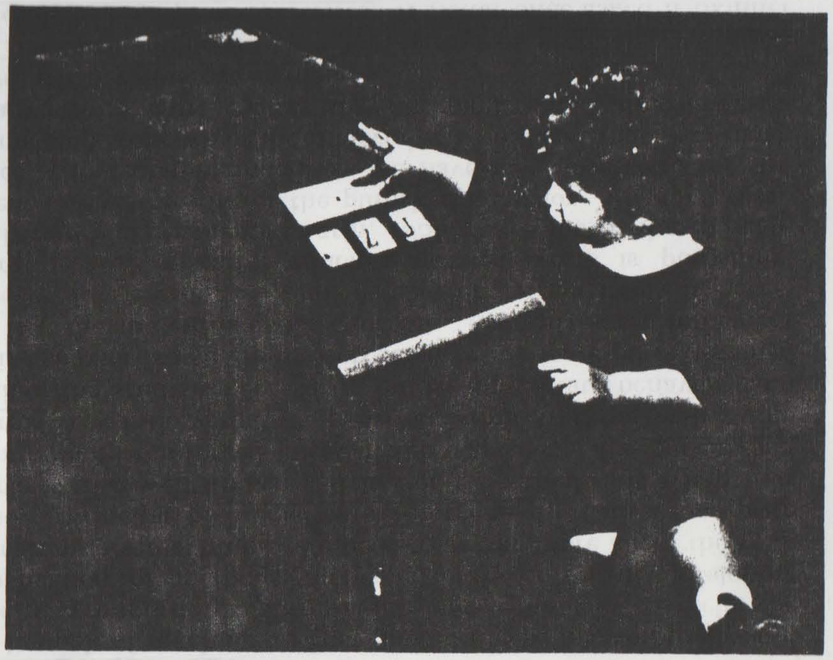


FIG. I. A selective-response machine for preschool children

as the sample, independent, in this case, of color or size. When the correct choice is made, the next frame is presented.

A teaching machine for a still lower organism is shown in Figure II.

This pigeon, with the aid of a teaching machine, has learned to hit the name plaque appropriate for a color projected above him. The principal difference between this and the other machines is that food reinforcement is used. With humans, simply being correct is sufficient reinforcement—pigeons will not work for such meager gains.

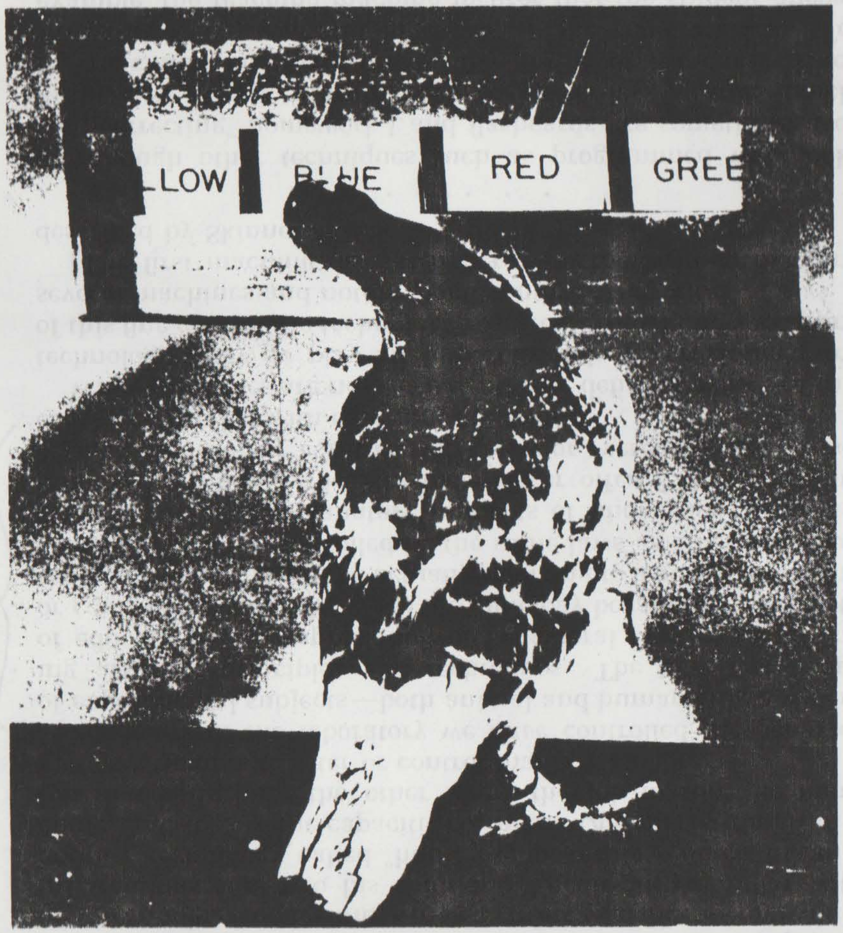


FIG. II. A teaching machine for lower organisms

Enough of machines. They should not be allowed to obscure the truly important feature of the new technology; namely, the application of methods for behavioral control in developing programs for teaching. We need to say no more about the well-known principle of immediate reinforcement. Our second principle is also well known. Behavior is learned only when it is *emitted* and reinforced. But in the classroom, the student performs very little verbally. However, while working with a machine, the student necessarily emits appropriate behavior, and this behavior is usually reinforced since the material is designed so that the student is usually correct. Not only is reinforcement needed for learning, a high density of correct items is necessary because material which generates errors is punishing. Laboratory experiments (Azrin, 1956) have shown that punishment lowers the rate of the punished behavior. In our experience with teaching machines we have also observed that students stop work when the material is so difficult that they make many errors. Furthermore, they become irritated, almost aggressive, when errors are made.

The third important principle is that of gradual progression to establish complex repertoires. A visitor once asked if Skinner had realized that pigeons were so smart before he began using them as subjects. The answer given by a helpful graduate student was that they weren't so smart before Skinner began using them. And indeed they weren't. The behavior developed in many experiments is like that developed in the classroom. Both are complex operants. Both require a careful program of gradual progression. We cannot wait for a student to describe the content of a psychology course and then reinforce the performance; nor can we wait for a pigeon to emit such an improbable bit of behavior as turning a circle, facing a disk on the wall, pecking it if lit, and then bending down to a new exposed food tray and eating. When developing a complex performance in a pigeon, we may first reinforce simply the behavior of approaching the food tray when it is presented with a loud click. Later the pigeon learns to peck a key which produces the click and the food tray. Still later, he may learn to peck this key only when it is lit, the peck being followed by the loud click and approach to the food tray. In the next step, he may learn to raise his head or hop from one foot to another, or walk a figure-8, in order to produce the lighted key which he then pecks; the click follows; and he approaches the food tray. This principle of gradual progression runs through many of the teaching machine techniques. Both human and avian scholars deserve the same careful tutorage. The teaching machine program moves in very finely graded steps, working from simple to an even higher level of complexity. Such a gradual development is illustrated in Table 1 by a few items taken from a psychology program.[2]

Table 1. ITEMS FROM THE PSYCHOLOGY PROGRAM

ITEM	CORRECT ANSWER	PERCENTAGE OF STUDENTS GIVING THE ANSWER
1. Performing animals are sometimes trained with "rewards." The behavior of a hungry animal can be "rewarded" with	Food	96
2. A technical term for "reward" is reinforcement. To "reward" an organism with food is to ________ it with food.	Reinforce	100
3. Technically speaking, a thirsty organism can be ________ with water.	Reinforced	100
50. A school teacher is likely, whenever possible, to dismiss a class when her students are rowdy because she has been ...... by elimination of the stimuli arising from a rowdy class.	Reinforced	92
51. The teacher who dismisses a class when it is rowdy causes the frequency of future rowdy behavior to (a) ......, since dismissal from class is probably a(n) (b) ...... for rowdy children.	(a) Increase (b) Reinforcement	86
54. If an airplane spotter never sees the kind of plane he is to spot, his frequency of scanning the sky (a) ...... In other words his "looking" behavior is (b)	(a) Decreases (b) Extinguished (or: Not reinforced)	94

The principle of gradual progression serves not simply to make the student correct as often as possible, but it is also the fastest way to develop a complex repertoire. In fact, a new complex operant may never appear except through separately reinforcing members of a graded series (Keller and Schoenfeld, 1950). Only this way can he quickly create a *new pattern* of behavior. The pigeon would not have learned the complex sequence necessary to receive the food if he had not learned each step in its proper order. Obviously a child can't begin with advanced mathematics, but neither can he begin with $2 \times 2 = 4$—even this is too complex and requires a gradual progression.

Our fourth principle is, in a sense, another form of gradual progression—one which involves the gradual withdrawal of stimulus support. This we shall call "fading." This method can be illustrated with some neuroanatomy material.[3] A fully labelled cross section of the medulla oblongata is placed before the student while he works with a large set of items pertaining to the spatial arrangement of the various structures. For example, "posterior to the cuneate nuclei are the ________." The answer is: "the cuneate fasciculi." After many such items, he begins another set and has another picture before him, but now the structures are only labeled with initials. A new set of items again asks a long series of questions pertaining to the spatial position of the various structures. For example, "between the gracile and the trigeminal nuclei are ________." The answer is the "cuneate nuclei." After many more items, he proceeds to a new set and the next picture. This time the picture is unlabeled. Again, he goes through a series of new items, not simple repetition of the previous ones, but pertaining to the same problem of the spatial location of the different structures. This set is followed by still another but with no picture at all. He is now able to discuss the spatial position of the structures without any visual representations of the structures before him. In a sense, he has his own private map of the medulla. He may further demonstrate his newly acquired ability by accurately drawing the medulla. The neuroanatomy example is an elaborate example of fading. Fading is also applied in a more simple form in constructing verbal programs without pictorial displays. A single item may in one sentence give a definition or a general law, and a second sentence in that same item may give an example in which a key word is omitted. This would be followed by a new example in the next frame but with the definition or law lacking.

---

[2] As of the time of writing, copies of this unpublished program, "A Self-Tutoring Introduction to a Science of Behavior," by J. G. Holland and B. F. Skinner, were available from the authors at Harvard University.

[3] This material was prepared by D. M. Brethower in collaboration with the author for use in research at Harvard University.

This brings us to our fifth principle: control of the student's observing and echoic behavior. In the classroom the student is often treated as though he were some kind of passive receiver of information that can absorb information spoken by the teacher, written on the blackboard, or presented by films. But all of these are effective only insofar as the student has some behavior with respect to the material. He must listen carefully or read carefully, thus engaging in usually covert echoic behavior. Ineffectiveness of classroom techniques is often credited to "inattention" or poor "concentration." It has been shown (Reid, 1953; Wyckoff, 1952) that if a discrimination is to be learned, adequate observing behavior must first be established. We have further found (Holland, 1958) that observing behavior—or, speaking loosely, "attention"—is subject to the same forms of control as other behavior. This control of observing behavior is of prime importance. When the student becomes very "inattentive" in the classroom, the teaching material flows on; but with a machine, he moves ahead only as he finishes an item. Lapses in active participation result in nothing more than the machine's sitting idle until the student continues. There is, however, a more subtle aspect to the control of observing behavior than this obvious mechanical one. In many of the examples we have seen, success in answering the problem depends only on the student's careful observation of the material in front of him at the moment. This may be illustrated by more material from the psychology program. A graph showing stimulus-generalization data is in front of the student while he works on the machine. In the program he may complete a statement like the following: "As the wave length changes in either direction from the wave length present during reinforcement, the number of responses   ." The answer is "decreases." The item serves only to control the behavior of observing the data. Of course, many more such items are used to discuss the same data.

This principle of controlled observation extends to the details of writing a single item; for example: "Two events may have a common effect. An operant reinforced with two reinforcers appropriate to different deprivations will vary with   deprivations." The answer is "two" or "both." Here the programmer's choice of the omission serves to insure a careful reading of the item. *Only* those parts of an item which must be read to complete correctly a blank can safely be assumed to be learned.

Our sixth principle deals with discrimination training. In learning spoken languages, for example, it is necessary to be able to identify the speech sounds. Students may listen to a pair of words on a special phonograph which repeats the passage as many times as he desires. The visual write-in machine instructs him to listen to a specific passage. For example, the student may hear two words such as: "sit, set." He listens as many times as he needs and then writes the phonetic symbols in the write-in machine, and operates the machine, thereby exposing the true answer and providing immediate reinforcement for his correct discrimination.

However, little academic education is *simple* discrimination. More often, it is abstraction or concept formation. An abstraction is a response to a single isolated property of a stimulus. Such a property cannot exist alone. Redness is an abstraction. Anything that is red has other properties as well—size, shape, position in space, to name a few. There are red balls, red cars, red walls. The term "red" applies to all of them but not to green balls, blue cards, or yellow walls. To establish an abstraction, we must provide many examples; each must have the common property, but among the various examples there must be a wide range of other properties (Hovland, 1952, 1953). This is best illustrated by examples from the preverbal machine shown in Figure III.

These are from a program which teaches a child to respond to the abstract property of form.[4] In each item, the upper figure

---

[4] This program was prepared by B. F. Skinner.

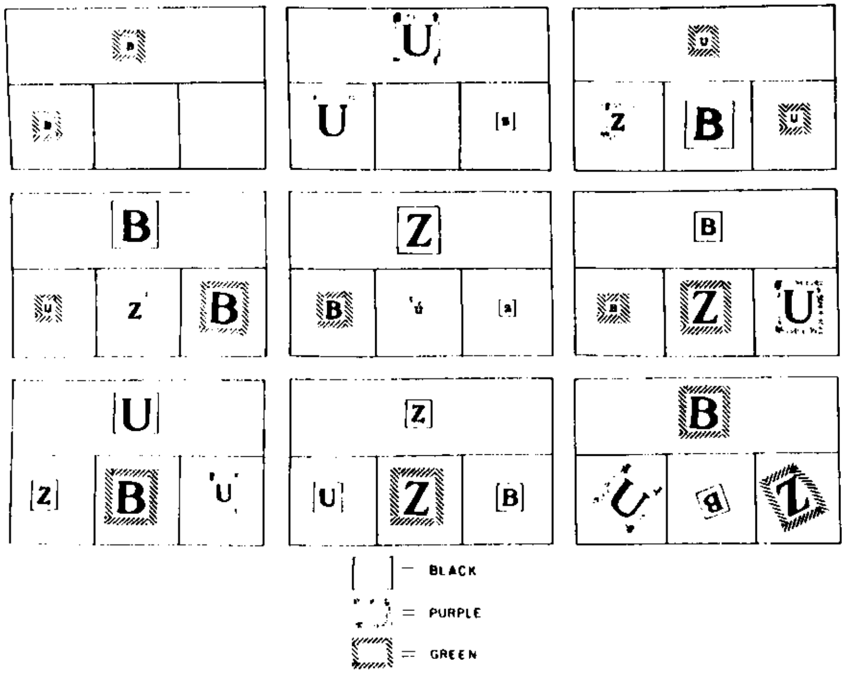


FIG. III. Selected items from a program which teaches young children to respond in terms of the abstract property of form.

is the sample and the lower three are the alternatives. While developing a program for establishing an abstraction, we remember our earlier principles and move through a gradual progression. The first several items would be like the first one; here there is a sample and a single match, the other two being blank. The sample and its match are exactly alike at this stage. After many such items, we would begin to have others like the next one in which the sample and its match again correspond in size, color and form; but an additional incorrect alternative has been added which differs from the sample in all these aspects. Later we move on to frames with three choices, again the sample and its match corresponding exactly. Next, the sample and the match may differ in some property such as color in the case of the next item shown or size in the next item. It is essential that the program contain many items among which the sample and correct match differ in all properties except the one providing the basis for the abstraction. Otherwise, the abstraction will be incomplete because the extraneous property will share some of the control over the abstract response. As we move on with additional examples, the sample and the correct match differ both in color and in size, and the incorrect alternatives are beginning to share some of the extraneous properties with the sample. The student continues with many such problems in which the only common property between the sample and the correct match is the shape, regardless of size and color. Even now our abstraction may be incomplete. We have kept the figures in only one orientation. Therefore, we also have a series in which the samples are rotated as in the next item. A great deal of academic education consists of trying to teach abstractions. Concepts such as force, reinforcement, supply and demand, freedom, and many, many other possible examples are all abstractions. Furthermore, in the academic setting, the student seldom adequately forms abstractions. The trigonometry student commonly uses triangles with the right angle as one of the two lower angles. If the triangle is rotated 90°, so that the right angle is upward, the student often does not recognize it as a right triangle. Neither is an abstraction developed simply by learning a definition. The psychology student who learns the definition of reinforcement in formal terms and is acquainted with a laboratory example of food reinforcement, may not realize the horrible consequences of sending his girl friend flowers to end an argument. Thus, in the psychology program, to develop a new concept, we follow the pattern in the preverbal example. A wide range of examples is analyzed; these examples differ in as many aspects as possible, but each still has the common property which characterizes the concept.

The last principle I shall discuss is really a question of a methodology which has served so well in the laboratory. This principle is to let the student write the program.

. . . .

When the student has trouble with part of a program, the programmer must correct this. The student's answers reveal ambiguities in items; they reveal gaps in the program and erroneous assumptions as to the student's background. The answers will show when the program is progressing too rapidly; when additional prompts are necessary; or when the programmer should try new techniques. When unexpected errors are made, they indicate deficiencies *not* in the student but in the program.

The most extensive experience with this principle of modifying the program to fit the student has been at Harvard (Holland, 1959) in the psychology program. In 1958 we had a program consisting of 48 disks or lessons of 29 frames each. After using the program and making a detailed, item-by-item analysis of the students' answers, we diagnosed the particular deficiencies in the program and revised it accordingly. The program was also extended to cover a larger amount of subject matter; and in 1959 it consisted of 60 disks. You have already seen a few items from the course (Table 1). After using the revised material in 1959, we evaluated the extent of its improvement. Table 2 shows the percentage of errors on the first 20 disks for each of the two years.

Table 2. A COMPARISON OF THE STUDENTS' ERRORS IN USING THE REVISED (1959) AND UNREVISED (1958) PROGRAM IN PSYCHOLOGY

	PERCENT ERRORS	PERCENT ITEMS IMPROPERLY SCORED BY STUDENTS
1958	20.1	3.6
1959	11.0	1.4

The revision eliminated about half the errors. The last column of the table gives percentage of improper self-scoring by the students. Revision also cut these scoring errors approximately in half. Furthermore, the revision decreased the time required to complete the material. Although the second year's material had more disks—60 as opposed to 48—it actually required the average student about one hour less to complete the work than it had required for the shorter first version. Frequency distributions on the median times in minutes for completion of the various disks are shown in Figure IV. These are the times required for the median student to move through each set of material answering every item once and to repeat the items answered incorrectly. Notice the considerable time required for

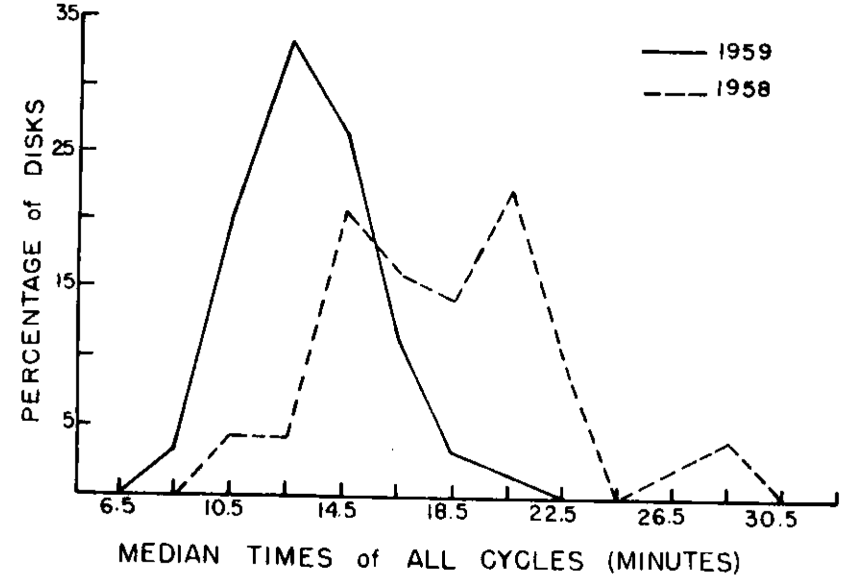


FIG. IV. Frequency distributions for the median times to complete the disks or "lessons" for revised (1959) and unrevised (1958) psychology programs.

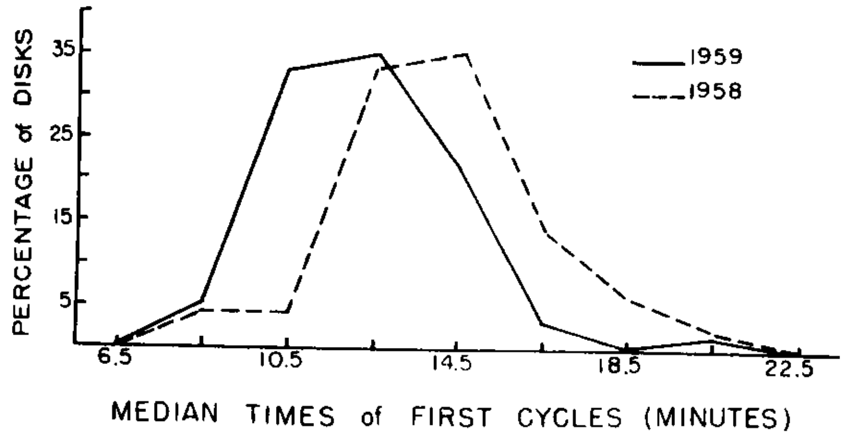


FIG. V. Frequency distributions for the median times to complete only the first cycles for revised (1959) and unrevised (1958) psychology programs.

many disks in the first year's material—primarily because students repeated the larger number of items missed in the first cycle.

But the improved material provided faster performance, even when the delay due to repetition of incorrectly answered items is not considered. The frequency distributions for the first cycle only are provided in Figure V.

These data exclude the time used in repeating items. Here too the revision produced slightly more rapid progress.

Such careful tailoring of material to fit the student is impossible with most teaching techniques. With teaching machines, as in no other teaching technique, the programmer is able to revise his material in view of the students' particular difficulties. The student can write the program; he cannot write the textbook.

We have seen that the principles evolved from the laboratory study of behavior have provided the possibility for the behavioral engineering of teaching. This new technology is thoroughly grounded in some of the better established facts of behavioral control. The future of education is bright if persons who prepare teaching machine programs appreciate this and appropriately educate themselves in a special, but truly *not* esoteric, discipline. But it is vital that we continue to apply these techniques in preparing programs. The ill-advised efforts of some of our friends who automatize their courses without adopting the new technology have an extremely good chance of burying the whole movement in an avalanche of teaching-machine tapes.

Please Note!

Report on the Initial Test of a Junior High-School Vocabulary Program

SUSAN R. MEYER

THE STUDY described here had two facets: It was a trial run of a part of a program and an investigation of elements in the design of teaching machines in so far as these elements could be duplicated in a programmed textbook. This programmed textbook contained 19 lessons, of which seven lessons were introductory in nature and 12 taught 26 common prefixes. The items were drawn from a larger program that teaches Latin and Greek elements which have come into English as roots and prefixes—a kind of "minimal repertoire" for vocabulary extension. The conditions provided by the teaching machine for immediate confirmation of each response and for the return to wrongly answered items were tested for their effects on the behavior within the learning sessions and on an outside criterion of learning.

I. OUTLINE OF THE EXPERIMENT

The programmed textbook, similar to that suggested by Homme and Glaser (1959b), was a "workbook" with the item on one page and the answer to the item on the following page. Next to the answer for Item 1 in any lesson was Item 2, with the answer to the latter being likewise next to Item 3 on the following page. Each student had his own textbook. He wrote his answer in blanks provided in the item and scored it by comparison with the correct answer on the next page. The content of a sample sequence of the program is shown in Table 1.

The outside criterion of learning was a prefix test, administered both as a pretest and as a post-test. The test had three parts. Part I listed 26 fairly common words and required the student to identify and define the prefix of each. Part II re-

---

This paper was prepared specifically for inclusion in the present volume. The work it reports was supported by grants from the Carnegie Corporation and the Ford Foundation.

quired the student to insert on a given root the prefix corresponding to the given definition. In both parts the words used were relatively common so that the student with a limited vocabulary would not be penalized. Part III was a test of "pure" prefix knowledge. An attempt was made by employing "nonsense" roots in the items to eliminate the effects of vocabulary size as far as possible. The student was given a definition of a "word" that could be constructed with a given nonsense root and the right prefix. The test was difficult, and its maximum score of 218 points was attained by no one.

Three eighth-grade classes in a junior high school in Concord, Massachusetts, worked through the materials. One class (14 students) was designated by the school as an honors group on the basis of its members' high ability and outstanding scholarship. In four one-hour English classes, these students did the first 10 lessons, alternating regularly between the programmed textbook and the Harvard teaching machine. They evaluated the two experiences and completed the program in the books with no further supervision. They had not been given the pretest. Each took the post-test upon finishing his book.

The three main experimental groups were drawn from two classes of superior readers. These eighth-graders had an average of tenth-grade scores in reading comprehension and vocabulary on a reading test (SRA). The two classes followed the same curriculum under the same teachers. On the basis of the reading-test scores and the students' scores on the pretest, these two classes were divided into three groups with equal means and . . . medians on both scores and approximately equal ranges.

The groups were treated as follows:

Group A: (16 students) The textbooks given to this group had no answers in them. Except for the absence of answers, the textbooks were identical to those given all the other students. The books were corrected by the experimenter after each class, and correct answers were inserted where the student erred or had failed to answer. The students were advised to look at their past performance before going on.

Group B: (15 students) These students scored themselves by putting an "X" on wrong answers as they proceeded through the lesson. The experimenter checked these books after each session for scoring accuracy.

Table 1. PART OF A PROGRAM IN HIGH-SCHOOL GRAMMAR

SENTENCE TO BE COMPLETED	TERM TO BE SUPPLIED
1. The prefix **circum** (think of "circle") means "around." The circumference of a circle goes the outside.	around
2. **Circumstances** are events going on you when something happens.	around
3. A person who sails around the world navigates the world.	circum
4. A cautious person **looks around** carefully before doing something; he is . . . . spect.	circum
5. Which of these airplanes would fly furthest? . . . . Least? . . . . (a) interstate, (b) circumglobal, (c) transcontinental	(b), (a)

Group C: (13 students) These students marked each item scored as "wrong" by putting a paper clip on the page as well as an "X" on the item. After completing the lesson, they went back through their books to each page so marked and wrote another answer. If the second answer was correct, they removed the paper clip. If not, they were to return again to the item after making the rest of the corrections.

Group H: (14 students) The honors class followed the procedure of Group B both on the machine and in their textbooks. Their performances were not checked during their learning sessions.

The procedure followed by Group C most closely replicated the procedure imposed by the teaching machine. The response to each item was followed by immediate confirmation. If the response was scored "incorrect," the student returned to the item at the end of the lesson. Group B did not repeat incorrectly answered items. Group A's responses were followed by delayed confirmation.

All work, including test-taking, was performed during regular English class periods under the supervision of the experimenter. The experiment covered 10 class days.

The following scores were available for most students: (a) vocabulary scores; (b) pretest scores (honors group excepted); (c) post-test scores and the "gain" score (pretest subtracted from post-test scores); (d) the error score in the textbook, ranked *inversely* to the number of errors in the correlations reported below; (e) the number of scoring errors, similarly ranked from low (good) to high (poor); (f) the number of errors made on each prefix; (g) the types of errors (spelling, omission, misreading, etc.) made and misscored.

II. SUMMARY OF RESULTS

The data were drawn almost exclusively from Lessons 8 through 19, which taught the 26 prefixes tested in the prefix test.

Learning Sessions

A. TIME TAKEN. The average total time for Lessons 8 through 19 was 132 minutes. The self-scoring groups, Groups B and C, reported slightly shorter times (129 minutes) than did Group A (137 minutes). Students wrote the time at which they started and completed each lesson on pages provided for that purpose, often with amusing results. Some finished a lesson two minutes after or even 45 minutes before they started it! Most lessons required eight to 15 minutes to complete according to the credible reports.

B. CHEATING. Since the answers were to be written on one page and the correct answer was given on the next, close supervision reduced the tendency to peek. However, one student managed, by shifting the page around, to read through the rather thin paper. After several warnings and in sympathetic response to very poor and punishing performance on honestly attempted lessons, this sixth-grade reader was "allowed" to proceed thus, and the data were discarded. Three other students, one extremely competent, erased wrong answers and wrote in correct ones. In general, the students performed honestly, taking in their stride the punishment that came with badly written or misplaced items.

Treatment Variables

A. IMMEDIATE CONFIRMATION. The achievement of Groups B and C was compared with that of Group A to test the effect of immediate self-scoring versus delayed "teacher" scoring. Absolute scores on the post-test were superior for Groups B and C, but $p = .06$. The amount learned (pretest score from post-test score) was significantly greater: $p = .03$. The difference in significance was due to the lack of equivalence in initial distribution on the pretest. Although the means and medians were about equal in the three groups, the quartile distributions were not. All of Group B's "below average" students were in the fourth quartile. The difference in the number of errors made in the textbooks between the self-scorers and Group A was highly significant: $p = .00005$. Students without answers in their books tended to lose the point of a long development and lacked the reminder of the subject of discussion provided by their seeing the expected answer. They also failed to answer far more often than did students who had to score the omission as an error. On the basis of evidence from misplaced items, presenting the correct answer does not appear to be an effective teaching technique; but it does produce more behavior (fewer omissions) and better behavior (fewer commission errors).

B. REPETITION OF MISSED ITEMS. The achievement of Group B was compared with that of Group C to test the effect of the correction of errors. No differences were found. The test used was the *Mann-Whitney U* test, and in every case, U_1 very nearly matched U_2. The two groups were equivalent in post-test scores, amount learned, and number of errors made in completing the program. The sole significant difference lay in the accuracy of self-scoring exhibited by the two groups. The added contingency of having to repeat items scored as "wrong" significantly increased the number of wrong answers allowed to pass as correct. Where such mistakes were "easier" to overlook, e.g., in the case of spelling errors, the differences were more significant. Five students in Group C and none in Group B failed to score an omission as an error.

The lack of difference between Groups B and C must be considered within the conditions of correction imposed by the way in which the textbooks were used here. A student who returned to a wrongly answered item had already written one answer which was still visible when the second answer was written, a condition that does not obtain on the machine. The program itself contained a good deal of "repetition," in the sense of many examples of each prefix in different words; thus students not repeating their errors had further opportunities to be correct and to be reinforced for any particular item they might have missed. Further evidence is needed under other conditions.

Self-Scoring by Eighth-Graders

Holland (1960) reported a percentage of 3.6 misscorings on the first version of his psychology program for Harvard students. The percentage was reduced to 1.4 in the revised version which had a comparably reduced percentage of errors. The percentage of error in scoring represented by the performance of the students in this study is 2.2, a comparable figure.[1] The data from this single version indicate also the same relationship between difficulty and misscoring as was found in Holland's two versions. In general, the greater the number of errors made on an item, the greater the percentage of misscorings on the item.

Investigation of individual data revealed some interesting details of scoring behavior: (a) It was not related to reading skills as measured by the SRA tests. ($R_s = .08$ for Groups B and C.) Members of the honors group, who were superior in reading skills to the other groups, were no different in accuracy from their equals-in-treatment, Group B. (b) It was not related to the pretest score as a measure of familiarity with the material. ($R_s = .30$, not significant for 28 students.) (c) Its relationship even to the number of errors made by each student ($R_s = .35$, $p = .05$) was no more than to be expected from the necessity of making errors in order to misscore them. Misscoring a correct answer "wrong" occurred infrequently and equally in both groups.

The percentage of misscoring wrongly answered items as "correct" was 8.8 percent of errors for Group B and 21 percent of errors for Group C. Such a wide variation in performance could be explained as a function of the punishing contingency under which each group operated. It is likely that scoring one's answer "wrong" is punishing under any circumstances, an observation which is supported by the reduction in the number of omissions under the self-scoring contingency. It is also likely that the repetition requirement makes scoring an answer "wrong" doubly punishing. Individual differences appear to exist in the "capacity for self-criticism." The program in its first major trial was sufficiently imperfect to make impossible an error-free performance. The added punishment of having to repeat errors could not, then, be avoided by the student's being doubly careful. There were no differences in the number of errors made. The avoidance of the double contingency was accomplished, rather, by "not seeing" the difference between one's own answer and the given correct answer. The data reported above are presented in Table 2.

Table 2.—SUMMARY OF DATA FROM PREFIX TEST AND TEXTBOOK

MEASURE*	GROUP				LEVEL OF SIGNIFICANCE		
	A	B	C	H	A—BC	B—C	B—H
Pretest (218 points) Range	60.5 (88-35)	62 (109-22)	61 (84-41)	—			
Post-Test Range	109 (135-77)	117 (156-81)	112 (141-83)	148 (172-126)	.06	N.D.	.001
Gain Range	35 (88-17)	56 (94-30)	52 (83-25)	— —	.03	N.D.	
Errors in text (481 items) Range	107 (42-211)	67 (36-196)	68 (14-86)	43 (24-97)	.00005	N.D.	.025
Omission errors Range	25 (7-68)	3 (0-48)	3 (0-15)	2 (0-10)	.00001	N.D.	N.D.
Misscore (+ and —) Range	—	6.5 (2-25)	14 (3-25)	7.5 (3-19)		.05	N.D.
Spelling misscore Range		2 (0-13)	5 (1-14)			.025	

* Figures reported are medians.

---

[1] This percentage represents the total number of misscorings in relation to the total number of answers by all students.

Predictors of Success

The findings agreed with those of others (Porter, 1959), in that measures of ability (here, the reading-test scores which would seem to be relevant) showed little relationship to the handling of the materials. The SRA scores correlated .25 with the pretest, .35 with the post-test, .34 with the error score, and .09 with the gain from pretest to post-test. These low correlations must be taken in relation to the classes involved. Although those few students whose reading skills were below the eighth-grade level showed the expected difficulty with the materials, there was almost no relationship above that level. Those students whose reading scores were below the tenth-grade class average *may* have been selected by the school for inclusion in the class because of "scholarship," while those students far above the average were not in the honors group for lack of it. The honors students were significantly superior on both error scores and post-test scores.

The pretest correlated —.52 with the "gain" score, indicating that, insofar as the test measured what was learned, students who knew little at the start of the experiment gained more from the materials than did the sophisticated students.

The greatest predictor of success on the post-test was the error score: $R_s = .69$ for groups A, B, and C combined; .82 for Group A; .82 for Group B; and .71 for Group C—the latter three significant at .01. The relationship between the total number of errors and the "gain" score, .49, was even more regular when broken down into the number of errors the individual made on those prefixes he learned and the number of errors on those he did not learn. Those prefixes eliciting few errors from the individual were those he learned.[2] The evidence supports the assertion that learning is more efficient if errors are reduced or prevented. Data are summarized in Table 3.

III. THE EFFECTIVENESS OF THE MAJOR PROMPTS USED

The term "prompt" is used here as a label for any technique employed to make emission of the correct answer more likely.

---

[2] One student made 11-percent error on those prefixes he gave correctly on the pretest, 11-percent error on those he learned, and 3-percent error on those he did not learn. However, in most cases, a 5-15 percent difference in the "right" direction was shown between learned and not-learned prefixes.

Table 3. CORRELATIONS AMONG THE SCORES FOR GROUPS A, B, AND C COMBINED

	PRETEST	POST-TEST	GAIN	ERRORS	MISSCORE
SRA Test	.25	.35 (p = .025)	.09	.34 (p = .025)	.08
Pretest		.43 (p = .005)	—.52 (p = .0005)	.16	.30
Errors*		.69 (p = .0005)	.49 (p = .0005)		.35 (p = .05)

* Ranked inversely to the number of errors.

The hoary educational technique of "repeat-after-me" was specifically excluded by Skinner in *Verbal Behavior* (1957e). More than one prompt was given in every item in this program. At no time did the student emit the "pure" pairs as given in the test. A pure item would be: "The prefix meaning 'around' is ______" or "The prefix circum- means ______."

In the program, blanks eliciting definitions were set in sentences, while the blanks eliciting prefixes were set on roots with which the prefix formed a legitimate word. In many items not all errors represented a violation of the prompt, and a double distinction had to be made between the effectiveness of the prompt in controlling writing behavior and its effectiveness in eliciting the correct answer.

The items in which the answer was available comprised the first three of the following types.

A. Copying in a Simple Item

Assuming no knowledge of the prefix or its definition, a standard format was adopted. The first example of such an item in Lesson 8 read: "The prefix *trans-* means 'across' or 'over' (from one place to another). A *transatlantic* flight goes ______ the Atlantic Ocean." In each item of this type the prefix was underlined, the definition was in quotation marks, and the student was required to copy one or the other in relation to a fairly common word. Other "simple" copying items specified the prefix to be used, leaving the student to spell it correctly. In these items a student employing another prefix violated the prompt, while a student misspelling the correct prefix did not. The latter errors were not counted in this tabulation.

There were 45 simple copying items in Lessons 8 - 19. No errors were made on 13 of these items, few errors on the rest. The 73 errors elicited by the 45 items represented 3 percent of the total responses. Of the 73 errors, seven were copying of the wrong member (prefix instead of definition, or vice versa), and 15 represented misspellings of the prefix to be copied (not to be confused with misspellings of a prefix given to be spelled correctly). Although these errors represented poor reading of the item, the five omissions and the 46 remaining errors would seem to indicate that these items had not been read by the students.

Types of errors are illustrated by an item in Lesson 9. The standard presentation of the prefix and its definition was followed by: "Interstate buses go back and forth ________ states." Four students wrote *inter-* in the blank, copying the wrong member of the pair. This result was undoubtedly facilitated by the too-close parallel between the two ends of the sentence. Two students wrote "*into*," having read the first sentence badly if at all and responding to the resemblance between this prefix and the previously learned *in-*. One student wrote "*from*"; he was obviously not under the control of the textual stimulus and somewhat dubiously under the control of the semantic requirements of the sentence, although a preposition was clearly required.

B. Copying in a Complex Item

"Complex" items were of three types: (a) two prefixes were given, and the student had to choose which answer belonged in which blank; (b) an example was given from which to construct a new example; (c) the answer was available in the text but was not isolated by underlining and quotation marks as it was in the "simple" copying items. One hundred and thirteen errors were made on eight items in this classification, an error rate of 24 percent. Twelve errors (7 percent) were committed on the three "choice" items. The two copying-from-example items produced 59 percent errors, a percentage of error which exceeded any other type of prompt. In both cases, over half of the wrong answers were not even the same prefix,

while the smaller group of wrong answers consisted of misspellings that the example was intended to preclude.

An illustration of copying-from-example is the following item from Lesson 14: "A flame *emits* or sends ____ light. A person who has left his native country is an ____migrant." The item was in a lesson dealing with spelling similarities of *ex-* and *dis-*, which were explained on a panel. Despite the strong hints to use the form of *ex-*, 22 students wrote *im-* and 17 students ignored the example to the extent of writing *em-*. (This was not the only word that students refused to believe existed in the language!) Owing to the lesser selective strength of the root as a prompt, the next "example" item drew a greater variety of responses. An item pairing *retain* with *abstain*, which was unsuccessful in eliciting *abs-* as the prefix of the latter (20 of 42 errors were "*ab-*"), was followed by: "If 'keep away from' is *abstain*, you could guess that 'take away from' would be ___tract." Of 11 "*ab-*" responses, only two were in self-scoring groups; 10 responses were *re-* (reverberation from the preceding pair?); two were "*sub-*"; two were "*dis-*"; three were "*ex-*"; and one was "*de-*"—a total of 25 (50 percent) errors.

The above examples, as is immediately evident, do not indicate that eighth-graders cannot use an example if instructed to do so. Such instructions were not given. All that can be deduced is that these students did not instruct themselves to use an example.

C. Copying from a Panel

Eight items in the program required a specific word from the panel accompanying the lesson. The sentence containing the blank was followed in each case by "(See panel.)." The panels generally contained several "rules" or a paragraph of explanatory material. Had the responses been called wrong only if a word *not* on the panel was supplied, the percent of error would have been considerably lower. Items of this type were classed as "complex" copying because a search of a large amount of material was required. Thirty percent of the responses were incorrect. Of the 134 errors, 25 were omissions.

Although most of the data discussed here were drawn from the later sections of the material, the most revealing panel item was the first such item in Lesson 1. The panel presented definitions of prefixes, suffixes, and roots which were practiced in the lesson. The three words were underlined and were given in the plural. The item asked the student to find the "name" of the piece *-er* on "*harder.*" Of the 15 errors, two were "*prefix*" and one "*root.*" In the given sentence: "The -er is a ...," eight students wrote "*suffixes,*" one wrote "*sufixes,*" two wrote "*suffixe,*" and one wrote "*suffex.*" ("*Sufixes*" and "*preffixes*" turned up throughout Lesson 1.) Toward the end of the lesson, most students were spelling the words correctly, although they misscored their early approximations.

The responses to the panels for Lessons 13 and 17, which presented "a basic idea" about words containing the prefixes discussed (*con-* and *ad-*), were uniformly poor, although the phrase "the idea of" occurred both in the panel and in front of the blank to be filled in. Just under half the students failed to line up the phrases on the panel and in the item that paralleled each other in order to copy the correct word.

D. Indications of the Number of Letters Or Words To Be Supplied

This type of prompt was employed in 80 items. Almost all correct answers in the program were single words or prefixes. In three cases a multiple-word answer was specified by blanks such as . In 77 items the prefix answer was specified by blanks indicating the number of letters such as " root" or " root." Few violations of the prompt occurred. Only 59 errors, or 1 percent of the responses, were shorter or longer answers than those specified by the blanks. The total percent of error on the items was far higher, 18 percent. Many of the errors represented some amusing violations of English spelling habits as students tried to fit their chosen answers into the spaces provided. For instance the blanks " spect" can be filled by only one prefix if an English word is to result. One student wrote "*tran-*," violating the prompt and the recently learned spelling of that prefix. Examples of errors "induced" by the three-letter requirement were "*sus-*," "*exs-*," and "*asp-*." Although "*asspect*" turned up elsewhere, "*aspspect*" occurred only in this instance.

In Lesson 16 four consecutive items introduced a set of *-erior* words, all of which have the prefixes taught in the program: *anterior, posterior, superior, exterior, interior.* Each

item had two blanks to be filled; and all blanks indicated the number of letters to be written, as follows: "The front part of something is called the rior. The middle part (in between) is called the ior." The items constituted a review of *super-, inter-, post-,* and *ex-* in the middle of the lesson on *ante-*. The results were remarkable. Of the 138 errors on the seven blanks, only two were violations of the prompt. In an item contrasting the "front" and the "back," the first in the sequence, the response "*a n t e*rior" was successfully elicited, since this was the prefix being taught in the section. The blanks "p erior," however, elicited: four—"pr *o t*," five—"pr *o s*" (the prefix *pro-* had been recently reviewed), four—"pa *n t*" (this is related to the "opposition" items discussed below), two—"pa *s t*," and a variety of less recognizable sequences. The antonym ***inferior*** given (not asked for) to elicit *superior* produced three answers of "*e x f e* rior," a "reasonable" deduction, but also a "*c o f e* rior" and a "*c o n f e*ior." The most original response was given by a student in Group B to the first occurrence of "the middle (in between)" mentioned above: "*m i d d l*ior." Another wrote "*o u*terior" for "'*exterior.*" The unusual responses elicited by these blanks were by no means confined to the less capable readers in the classes. As in the discussion of responses to "*prospect,*" it is also evident here that this type of prompt did not produce the desired "self-editing." Instead of searching further for a word or prefix that satisfied the requirement, the students tended to operate in a Procrustean manner on their prepotent responses.

E. The Root As a Formal Prompt

Giving the second half of a word limited the number of choices for the first half that a reader of English would recognize as correct. In order to recognize the correctness, of course, the student must have the whole word in his vocabulary. If the completed word is not in his vocabulary, the root cannot act as a prompt. Using the Thorndike word frequencies as a measure, the roots supplied for completion were divided into three groups: (a) those which, with the correct prefix, produced an AA or A word (first- to third-grade vocabulary); (b) those which produced a word with a frequency between A and 20 (fourth-grade vocabulary); and (c) those which were ranked below 6 (high-school vocabulary and beyond). The percentages of error were: 10 percent for the AA-A words, 12.5 percent for the fourth-grade words, and 22 percent for words above the eighth-grade level. The difference between the percent of error on unfamiliar words and that on the most familiar can be ascribed to the strength of the relationship (although backward) between the two parts of the word which made the correct response more likely when the result was familiar to the student. However, in working with many unfamiliar words, a recognizable product may have become less of a primary goal for the students. Of 18 errors made on *dif*fident, a word not considered to be in the students' vocabularies, eight were the prefix spelled *"dis-."* Of the 32 errors made on *dif*fer, 12 were the prefix spelled *"dis-."* A further illustration is an item in Lesson 19 containing a "complex" copying prompt: "The prefix *sub-* assimilates with *c.* A *ceed* word meaning 'come next' (after or 'under') is ...... ceed." The response *"sub-"* was given by 16 students, three of whom misscored their answer. One wrote *"su-"* and one wrote *"pro-,"* thus violating the copying prompt but not the root-prompt as had the misspellers.

It is difficult to extract evidence for the "strength" of a root as a prompt in relation to the number of different prefixes it might elicit. The case of " migrant" strongly controlling the known *im-* rather than the unknown *e* has been mentioned. The roots *-tract* and *-spect* tended to produce a variety of erroneous but plausible responses when the desired prefix was weakly determined, as in review sections. The strength of a root, however, was a function of more than the number of *true* words it could govern. In a sequence teaching the prefix *per-* in Lesson 18, the root-suffix combination *-spicuous* might be expected to elicit the more frequent *con-* rather than the unknown *per-*, which depended on inter-item sequence alone. It did elicit three *con-* responses, but 13 errors were *"sus-"* or *"su-."* An item in Lesson 15 gave the root *-gregate*, the correct prefix for the definition given being *con-*. The response *"seg-"* occurred twice, *"se-"* not at all. Two students wrote *"inte-"* (*"integregate"*), and one wrote *"inter-."* One student caught enough of this error to respell the root with the result: *"intergrate."*

A part of a word given as a prompt shares with the indication of the number of elements the requirement that the students have a high level of sophistication in the structure of English words. Where the word to be constructed is unfamiliar, editing his response would be limited to probable letter and sound sequences in the language. This sophistication need not be verbalized; a student may "feel" something is wrong without "knowing" that an *s* never follows an *x* or that an *n* almost never precedes a *p*. A student writing *"aspspect"* or *"exsspect"* might be expected to question the product, as might a student writing *"disfer"* or *"subceed."* That superior eighth-graders did not respond to the absurdities they created is readily evident in the data.

F. Inter-Item Prompts

The prefix-definition pairs were typically introduced in an item requiring the student to copy one or the other (the "simple copying" item discussed above). The item immediately following it required either the same member of the pair or the noncopied member. In both cases the student had presumably just read both, although he had copied only one. The percentage of error was less on items requiring the same response than it was on items requiring the other member. Of 15 items requiring the same member, six elicited no errors, eight elicited six or fewer errors, and one elicited 15 errors. The average error was 5 percent. Of seven items requiring the other member of the pair, none was perfect. The average was 11-percent error. The single high-error item in the "same-member" items was distinguishable from the other second items on two counts: the copying item preceding it introduced two definitions, and the root given in the item produced a word unfamiliar to many of the students: *perverted*. The definition (a sense of humor that is "turned completely" upside down) also fitted the most common erroneous response. Seven students wrote *"in*verted." Although the correct prefix *"per-"* had been written twice in the preceding copying item, five students wrote *"pre*verted"; one *"a*verted" and two *"con*verted" brought the total of errors to 15.

In the construction of the program, sequence was a major prompt. Students were introduced to a prefix and its definition and then were led through a sequence of items in which the same subject was discussed and in which their responses would be correct if they did not deviate from it. That it was effective was evident from the number of students who correctly constructed *circumspect*, *perspicuous*, *antecedent*, and other words with which they were not likely to be familiar. Several students even noted with indignation that they had "never heard of" such words. The frequent use of this prompt had unexpected effects, however. After several items in which *anti-* was put in front of whole words (such as *Antinazi*), an item in Lesson 16, asking for the opposite of *sympathy*, produced 16 answers of "*antisympathy*" or "*antisimpathy.*" The prompt was not, however, strong enough to keep the students in Group A on the track. In Lesson 14, which collected the greatest number of errors from all students, one member of the no-answer group missed the last 19 items, and nine others missed 10 or more of the last 15. Where other strong prompts were working against sequence, such as a more common word, the tendency to ignore the "subject" was evident in all groups. The self-correcting groups were generally put back on the track by seeing the expected answers.

G. Minor Prompts

Almost all items requiring the student to emit the definition of a prefix used English sentence structure as a prompt. In filling a blank such as: "A subway goes   the regular roadways," the student was limited to a small group of words, prepositions which generally occupy such a space in a sentence. The prompt restricted the choice of answers, but it could not be said to determine the answer. There were few violations of English syntax in the many examples of this type of prompt.

In three items in which a possible response was pre-empted, e.g., "A person who is chosen out of a group is   lected (or selected)," the pre-empted response was given five times. In the above example, a student who wrote "*se-*" instead of "*e-*" in the blank violated the prompt and also violated the semantic requirements of the whole sentence.

A few items in the program were constructed in a manner intended to contrast two prefixes. In some cases the prefixes were opposites (e.g., *in-* and *ex-*); in others a difference in meaning of two words was in question (e.g., *export* and *deport*). One such item read: "If you 'go against' a law, you   vene it; a group that is 'going together' is   vening." It occurred in the middle of a sequence on the prefix "*contra-*" and constituted a review of "*con-*." The item elicited 26 pairs (the same response in both blanks), of which six were pairs incorrect in both blanks, three of the latter being nonexistent. Seven students wrote "*con-*" in both blanks, 13 students wrote "*contra-*" in both blanks, three wrote "*inter-*," one wrote "*de-*," and one wrote "*pre-*," and one wrote "*intra-*" in both blanks. The most successful item of this type explicitly stated the opposition involved; in the middle of a lesson on *dis-*, this sentence appeared: "The opposite of consent is   sent." Only four students wrote "*con-*" in the blank! On 10 such items the "same" response was given 24 percent of the time. Wrong but dissimilar prefixes increased the error scores on items of this type to which the responses were uniformly poor.

Data on prompts are given in Table 4.

Summary of Effective Prompts

Presenting a sentence, or "rule," which defined the prefix that was being taught proved to be a successful method of introducing the prefix-definition pairs. Students were generally successful in copying into the example the required member of the pair and in performing it correctly on the next item in the sequence. The percentage of errors increased when the immediately following item required the noncopied member of the

Table 4. SUMMARY OF PERCENT OF ERRORS REPRESENTING VIOLATIONS OF THE MAJOR PROMPTS

TYPE OF PROMPT	NUMBER OF ITEMS	PERCENT ERROR
"Simple" copying	45	3%
"Complex" copying	8	24
Copying from panel	8	30
Inter-item sequence: same response	15	5
Inter-item sequence: other response	7	11
Restriction on number of elements	80	1
Root half of common word (AA-A)	18	10
Root half of common word (fourth grade)	26	12.5
(High-school vocabulary	85	22)*

*Included for comparative purposes; not a prompt.

pair. A large increase in errors was found when the correct response was available but not highlighted in the item by such indices as quotation marks or underlining. In like manner, students showed considerable difficulty in locating the suitable unemphasized word in the panel material accompanying the lesson.

The responses given by the students indicated that readers at the level represented by the three classes used in this study were sensitive to sentence structure as a determinant of the class of words belonging in the answer space. They were less responsive to factors determining the similarity or difference of two answers in items requiring such discriminations. The misspellings, both emitted and misscored, indicated an insensitivity to probable and possible letter sequences within English words. In linguistic terms, "syntactic" prompts were effective; "semantic" prompts were less so; and "orthographic" prompts were ineffective. The latter could hardly be said to be prompts. The regularities that do exist in English need to be taught before they can be used to teach new material.

From the first year's data presented here, a set of "rules" could be deduced for successfully introducing a set of "paired associates" such as these prefix-definition pairs. There is sufficient evidence in the students' responses to indicate that copying an answer from a clearly designated "rule" is not an unnecessarily elementary requirement. A prefix or definition that has been copied has been seen, although it may not be permanently learned. Once copied, it is available for use on the next item. The other half of the pair is not necessarily available and could likewise be introduced by a copying item. "Irrelevant" prompts such as type size or even type color might be worth trying for emphasis on key words in larger texts such as panels. Other words as models for new words might be clearly designated as such. The students have not only suggested by their errors many items that need revision, they have also pointed to certain techniques of presenting material that likewise need revision.

Note

An Exploratory Investigation of an Apparatus for Studying and Teaching Visual Discrimination, Using Preschool Children

WELLS HIVELY

In a common educational task, particularly in the early grades, children are taught to classify objects, pictures, and symbols by selecting or sorting them in response to various instructions. Psychological experiments concerned with "discrimination learning" or "concept formation" often involve a similar process—a simple response is differentially reinforced with respect to certain properties of a sequence of stimulus displays.

In discrimination experiments an apparatus which automatically presents stimuli and records responses is usually more accurate and efficient than a human experimenter. When used in the classroom such a device is, of course, a "teaching machine" (Skinner, 1958c). The following is a report of exploratory research designed to see whether a practical apparatus can be developed for studying and teaching visual discrimination in children automatically.

The apparatus displays stimuli beneath small windows. Subjects respond to the stimuli by touching the windows, which are arranged so that a light pressure will close an electrical contact to operate automatic programming and recording equipment. Various arrangements of windows permit the study of different kinds of stimulus control—the arrangement studied in the present investigation was "matching to sample."

APPARATUS

Three plexiglass windows, each measuring 2¼ by 3¼ inches, are spaced ¼ inch apart in a triangular display (Figure 1). A

This paper is published for the first time in this volume. This research was supported by a grant from the Carnegie Corporation. It was done under the direction of B. F. Skinner and profits from the advice of James Holland, Douglas Porter, and others associated with teaching machine research at Harvard University over the past three years. The interest and assistance of Judith Schoellkopf of the Harvard Preschool, Mrs. Carl Friberg of the Cambridge Nursery School, and Paul Poehler of the Belmont Day Camp are gratefully acknowledged.

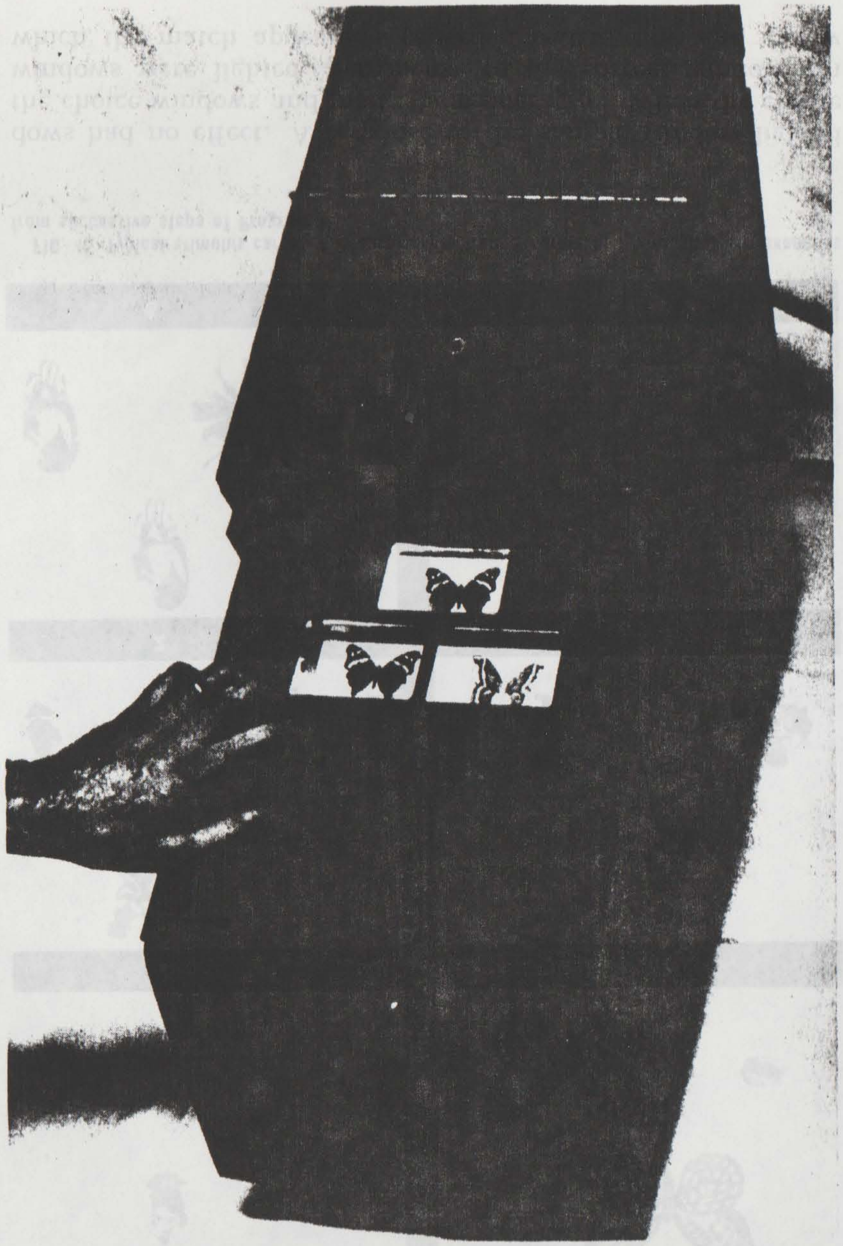


FIG. I. A two-choice preverbal machine for teaching visual discriminations

series of 5-by-8-inch cards is automatically presented beneath the windows, each card carrying three stimuli: a "sample," which appears in the top window, a "match," appearing in one of the two lower "choice windows," and an "alternative," appearing in the other. Each of these stimuli may be separately illuminated. Each window is hinged, spring-loaded, and provided with an electrical contact, and each card has coded holes cut along its edge which actuate programming contacts.

The display panel and card changer make up a unit approximately one foot wide, two feet tall, and two feet long, which rests on the floor and may be operated conveniently by a subject either standing or seated in a child's chair. Programming and recording circuitry are housed in a remote unit for convenient observation, operation, and shielding from the subject.

STIMULUS MATERIAL

Stimuli were pictures cut from a variety of children's books ("sample" and "match" being cut from two copies of the same book). Two series of stimulus cards, or "programs," were made up from them.

In Program I each card carried only two pictures—a "sample" and a "match." The "alternative" was always the plain white surface of the cards (Figure II-A). Pairs of pictures in this series varied widely in size, color, and content but occurred in a random order.

In Program II each card carried three pictures. The program began with cards on which the matching pictures were much more salient (as judged by size and contrast) than the alternative (Figure II-1) and progressed gradually to cards on which the alternative was much more salient (Figure II-5).

In each program, "matches" appeared in each choice window an equal number of times but in a random sequence. Each program included approximately 120 cards, divided into three equal sets.

A Test Set included 40 cards sampled from Program II and arranged in random order.

OPERATION OF THE APPARATUS

Each time a stimulus card was presented, only the sample window was lighted. Responses to the unlighted choice win-

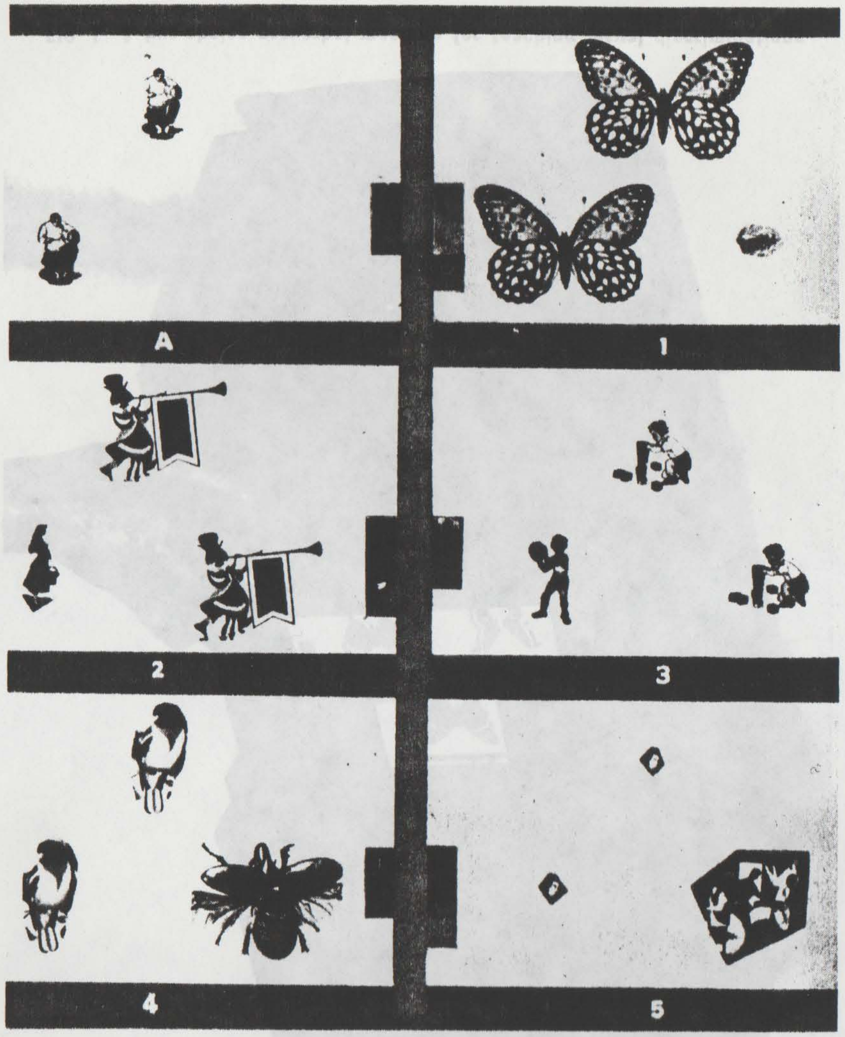


FIG. II. Typical stimulus cards. A is an example from Program I; 1 through 5 are examples from successive steps of Program II.

dows had no effect. A response to the sample window lighted the choice windows and made them operative. When the choice windows were lighted, a response to the correct window (in which the match appeared) caused a bell to ring and a new stimulus card to be presented. (Changing cards required approximately 3.75 seconds, during which all windows were dark.) A response to the incorrect window (in which the alternative appeared) darkened the choice windows and made them inoperative. Another response to the sample window was required in order to light and activate the choice windows, another choice response could then be made, and so on.

SUBJECTS AND PROCEDURE

Twenty-seven children, ranging in age from three to five and one-half years, were studied in nursery schools. The apparatus was placed in a corner of the playroom with the control unit about eight feet away. A subject was usually chosen by the nursery-school teacher and brought to the experimenter; the teacher then withdrew or remained silent while instructions were given by the experimenter (*E*). Since the main purpose of the research was to see how effectively the apparatus *alone* shaped and maintained the behavior of the children, the instructions were limited to the following:

1. "Would you like to play with this game?"
2. "You can make it go like this." (*E* operates machine once.)
3. "Now you can do it." (If *S* does not respond,[1] ceases to respond, or asks questions, *E* repeats Instructions 2 and 3.)
4. When *S* has responded to three stimulus cards, the experimenter says, "You can play with it as long as you want to," and withdraws to the control unit behind the subject. Other questions are answered by, "You can play with it as long as you want to."
5. The session is terminated when *S* finishes the set of cards, leaves the room, or stops responding for two minutes. "All finished? Good, you can play again another time."

All subjects were given Program I. Those who reached a performance criterion were then given Program II, ending with the *

---

[1] The abbreviation *S* is used to denote the "subject" (student) in the experiment.

Test Set. Each subject worked on one set per session, and sessions were separated by a day or more.

A polygraph recorded responses and card changes, permitting observation of errors, rates, and patterns of responding.

RESULTS

One of the purposes of this investigation was to see whether the apparatus provided enough reinforcement to keep children working in competition with other preschool activities. Within the modest limits which were set, it did.

The time taken to work through a set of 40 cards ranged from four to 13 minutes. Twenty-six of the 27 subjects completed each set during a single session. One subject stopped occasionally before reaching the end of a set but went to work again each time his next turn came. Twenty-five subjects returned for as many sessions as were offered; two declined to continue after the end of the second session.

Further investigation is needed to determine how much the operation of the mechanism, the type of stimulus material used, and the social variables involved in the experiment contribute to this motivational effect.

The subjects' performances are summarized in Table 1. Only the *first* response to the incorrect choice window for a given stimulus card was counted as an error. A "percent-error" score was calculated by dividing the total number of errors by the number of cards (50-percent error is therefore a chance performance). The performance criterion was 10-percent error or less for one complete set of cards.

Program I required the subjects to learn a simple discrimination. Fifteen of the 27 did so. In Program II the stimulus control was gradually altered until matching was required. As measured by performance on the Test Set, four out of 13 subjects learned the matching task. Clearly, the apparatus did not teach these discriminations efficiently. However, an examination of the records of the subjects who did not learn the required discriminations showed that most of them were learning something else—behavior which was shaped and maintained by "accidental" operation of the apparatus.

Table 1. SUMMARY OF PERFORMANCES

			ERRORS MADE BY SUBJECTS WHO REACHED CRITERION			
SUBJECT	SEX	AGE	PROGRAM I	PROGRAM II	TOTAL	CONFLICTING BEHAVIOR
1	F	3-3	0	—		S
2	M	4-0	2	5	7	
3	M	4-10	4	—		S*
4	F	5-6	5	11	16	
5	F	5-1	6	31	37	
6	F	5-2	7	(Subject unavailable for Program II)		
7	M	3-0	7	(Subject unavailable for Program II)		
8	F	4-6	8	—		C
9	F	4-11	13	—		S*
10	F	3-7	13	—		S
11	F	5-6	18	—		S
12	F	4-1	19	—		
13	M	3-11	22	—		
14	F	4-6	23	39	62	
15	F	5-0	36	—		C
16	F	3-5	—			
17	M	4-1	—			P*
18	M	3-10	—			P*
19	M	3-5	—			P
20	M	3-9	—			S
21	M	4-5	—			C
22	M	4-8	—			C
23	F	3-5	—			C
24	F	3-6	—			C
25	F	5-0	—			C
26	F	3-9	—			C
27	F	3-10	—			C

S = Simultaneous responding.
C = Chained responding.
P = Systematic responding in relation to the position of the choice windows.
* Partial control by the stimulus material.

During Program I, three subjects responded systematically either to a single choice window, to the window which last operated the apparatus, or to the opposite window from that which last operated the apparatus. The records of two of these subjects showed an interaction of this systematic responding with respect to position with partial control by the stimulus material.

The apparatus was arranged to record an error if the two choice windows were pressed at once (within 1/10 second of each other). This requirement was not enough to prevent "pushing-both-windows-at-once" from becoming a stable form of response for one subject during Program I and for three during Program II. The records of two others during Program II showed an interaction of this "simultaneous responding" with partial control by the stimulus material.

As described above under "Operation of the Apparatus," responses to unlighted choice windows had no effect, responses to the sample window while the choice windows were lighted were ineffective, and all responses were ineffective while the cards were changing. But since the subjects responded freely, they sometimes included these ineffective responses in a sequence which happened to operate the apparatus, and these response sequences were strengthened and maintained as a whole. Seven subjects during Program I and two during Program II developed this "chained responding," which ranged from long variable sequences to short relatively stable chains.

DISCUSSION

A large part of the behavior of subjects who did not come under the control of the stimulus material can be explained as the result of accidental contingencies of reinforcement. A number of variables entered into these spurious contingencies.

Simultaneous responding was strengthened when an incorrect response occurred together with or shortly after a correct response. Operation of the apparatus was made contingent upon holding the correct choice window down for a given interval while not pressing the incorrect window. The 1/10-second interval used did not prevent simultaneous responding—other values of this interval need to be studied.

When an incorrect response occurred shortly before a correct response, they tended to be strengthened as a chain. The apparatus may be arranged to operate only when a response to the correct window is not preceded by a response to the incorrect window for a given interval. Various durations of this interval, with and without correlated stimuli, need to be studied.

An alternative to requiring that incorrect responses not be followed by correct ones is to require that correct responses always be closely preceded by other correct ones—by making it necessary to press the correct window a given number of times in succession in order to operate the apparatus. The effect of requiring various "ratios" of correct responses should be studied. There are indications that this may be a very effective variable.

Chance correlations of a variety of stimuli with those upon which reinforcement was contingent seemed to account for some of the behavior which interfered with the development of the discriminations. Systematic responding with respect to the window positions is an example of this effect. Operations which increase the relative importance of the critical stimuli should hasten the development of control. (This consideration suggested the use of a window-press manipulandum, which brings the locus of response as close as possible to the stimulus material.) The use of only two choices probably maximizes distracting position cues; larger numbers of choices certainly need to be studied. Other important variables include the illumination of the stimuli in relation to the surrounding area and the location of the windows in the display panel.

Operations which sharpen contingencies necessarily reduce the frequency of reinforcement. Naive subjects may extinguish before coming under control of what proves to be the most effective set of variables. This problem may be solved by gradually making contingencies of reinforcement precise while maintaining a given frequency of reinforcement. For most practical purposes, instructions may also solve the problem, but it should be emphasized that instructions are probably not a *substitute* for precise contingencies of reinforcement. Although one may see to it that subjects are responding correctly on one set of material, there is no guarantee that a change in material may not result in changes in behavior which allow unwanted contingencies (if they exist) to take hold. In the present experiment the gradual stimulus changes in Program II had this effect on five out of 13 subjects, although all had achieved an accurate performance on Program I.

SUMMARY

An apparatus for studying and teaching visual discrimination automatically is described, together with the performances of 27 preschool children who operated it in exploratory experiments.

Simple discrimination and matching performances were taught with moderate success. A number of factors which need further study are suggested—most of which involve the elimination of spurious contingencies of reinforcement. The prospect of developing an efficient "associative discrimination teaching machine" seems good, and such a device has exciting possibilities as a means of teaching and testing the very young, the handicapped, and the mentally deficient. It also offers the possibility of studying human discrimination with a degree of experimental control comparable to that obtainable in experimentation with animals.

PART IV

Contributions from Military and Other Sources

THE PAPERS in Parts II and III of this volume represent in each case relatively unified lines of effort initiated, respectively, by Pressey and Skinner. The papers in Part IV reflect more diverse origins, including work conducted under United States military programs, and contributions from scientists in industry both in this country and in England.

The initial paper, by Ibert Mellan, is unique in that it deals with the history of educational inventions by identifying patents for educational devices. These are patents issued by the United States Patent Office from the early part of the nineteenth century up to the early 1930's. A limitation of this paper is that the patents are, for the most part, merely listed rather than described. By no means all of the devices are self-contained "teaching machines" that provide immediate feedback in a sequence of instructional items. Nevertheless, the document is interesting as a record of man's inventiveness, in one country at least, in attempting to improve educational effectiveness through the use of mechanical teaching aids. The record presented by Mellan may, also, be of some value as a source document to provide a point of departure for more detailed examination of the surprisingly large number of such inventions which were patented even prior to the development of the first Pressey machine in the early 1920's. Unfortunately, no more recent listing has been discovered to bring the record up to date for the past 25 years, and it was not possible to prepare a comprehensive compilation in the course of readying the present volume for publication. The editors have appended, following Mellan's article, a brief note on a few of the more recent patents which have come to their attention. A more systematic and complete examination of devices invented and patented since 1932 would be very useful. An analytic discussion of the features and rationale of various kinds of devices would also be of considerable interest.

The papers by Glaser and others, Briggs, Crowder, and Rothkopf deal with developments that stem from research and development projects conducted by the U. S. Air Force during the 1950's. Some of the devices they discuss have been briefly described in the paper by Lumsdaine (1959d) in the introductory section of this volume. The papers given here provide only a fragmentary picture of the quite extensive amount of attention that has been given by the military services to the development of instructional devices and techniques. Much of the relevant military work has centered around more general concepts of training devices and media and has not directly concentrated on the individual self-instructional techniques for presenting academic subject matter with which this volume is primarily concerned. Most of the focus of military effort has, rather, been on devices for the training and assessment of particular skills such as those required in the maintenance and, especially, troubleshooting of complex equipment systems. Another related area of work sponsored under military research programs has been concerned with the design and use of instructional films and related audio-visual media. Adequate coverage of the work in these two areas, though it has pertinence for the technology of teaching machines and self-instructional programs, could not be included in the present volume.

To a considerable extent the contribution to instructional technology resulting from military research and development reflects the participation of experimentally and behaviorally minded psychologists and educators in these efforts. This work is documented in the reports of several symposia sponsored by federal agencies. Among these are symposia held under the auspices of the National Academy of Sciences—National Research Council, and the Air Force—see Finch and Cameron, 1956 and 1958, and 1960 (in press)—as well as in other symposia sponsored by the Office of Naval Research and the Department of Defense.

A number of experimental studies of instructional films are quite relevant to the organization and presentation of subject matter in self-instructional programs. As noted by Lumsdaine (1959a), this is particularly true of studies in which alternate presentation and active student response to exercise questions have been employed. Brief abstracts of a few selected studies in this category are given in Appendix I (e.g., Hovland, Lumsdaine, and Sheffield; Kimble and Wulff; Maccoby and Sheffield; Cook and collaborators). These bear particularly on questions of the role of active response, size of demonstration and practice steps in procedural learning, and the use of prompts or response cues. For a more general description of military and civilian experimental research on instructional films the reader may be referred to summaries by Hoban and Van Ormer (1950), May and Lumsdaine (1958), Cook (1960), and Hoban (1960).

Some of the pertinent military research and development on training devices—particularly that dealing with trouble-shooting training—is also briefly abstracted in Appendix I. Among these studies is the series conducted at the University of Southern California, under contract with the Office of Naval Research, by Bryan and his associates; see also Cantor and Brown (1956) and other work described in the comprehensive compilation and review by Fattu (1956). Although not all of these studies deal directly with programmed-learning techniques, they have implications for self-instructional methods and form part of a common stream of work in which self-instructional features and other aspects of training equipment cannot always be cleanly separated from each other. Also of relevance in this connection is the work on military training problems conducted by the American Institute for Research—e.g., see abstracts in Appendix I of reports by Glaser and Glanzer (1958), by Klaus and Glaser (1958), by Lumsdaine (1960a,c), and by Miller (1953, 1954)—and the extensive in-service and contract research programs of the former Air Force Personnel and Training Research Center (AFPTRC) and its predecessor organizations. Abstracts of illustrative studies in the latter category include those by Briggs and associates (1956, 1957), Besnard and others (1955), Crowder (1957), and French and associates (1956, 1957).

Some of the military research and development has dealt more directly with individual self-instruction devices for teaching verbal behavior. Devices like the Navy's Self-Rater (mentioned by Skinner, 1958c) should be noted, as should the fact that a considerable amount of the work by Pressey and his students at The Ohio State University was sponsored by the Naval Training Device Center (formerly the Navy Special Devices Center). This Navy-sponsored work dealt with self-instructional techniques as such, and was not primarily oriented at the time to direct military training applications of a specific sort. The same was true of the sponsorship by the Army's Human Resources Research Office of some of Skinner's work on programming techniques see (abstracts of reports by Skinner [1957b,c,d] in Appendix I).

As examples of the military work on training devices which is more or less directly relevant to self-instructional techniques, several reports deriving from the program of the Air Force Personnel and Training Research Center (AFPTRC) are reproduced here in Part IV. The first of these, reported by Glaser, Damrin, and Gardner, was part of a group of studies conducted through a contract at the University of Illinois and is one illustration of the studies on trouble-shooting (see Fattu, 1956). This study, originating out of a concern with problems in measuring the proficiency of electronic maintenance mechanics, resulted in a technique which provided feedback during examinee testing. It thus afforded self-instructional properties somewhat in the same way as did Pressey's initial devices. Basically, the feedback technique devised by Glaser, Damrin, and Gardner (see also Glaser and Phillips, 1954) provided a variable sequence in which each response made by the subject yielded knowledge of results which was, in turn, used to direct his further responses.

Such techniques for proficiency assessment in situations involving variable or contingent sequences of responses lead rather directly to the variable-program self-instructional devices and to the technique of intrinsic programming, described by Crowder in the third paper of Part IV. The Tutor device described by Crowder is a current, commercially produced version of an earlier variable-program microfilm device which Crowder employed at the AFPTRC Maintenance Laboratory (see Figure VI on p. 297). The technique of intrinsic programming, devised by Crowder to achieve a variable-sequence program, is of considerable interest in that it represents a difference in rationale as compared with the approach followed by Skinner and by most of the other contributors to this volume. In Crowder's approach less emphasis is placed on prompting to insure that correct responses are evoked during learning; rather, the student's responses serve primarily a "quiz" function, to determine whether he has learned from reading the material preceding each question. His response to each question then determines what material he is exposed to next, the latter either carrying him on further in the development of the subject matter or shunting him off to correct a misconception revealed by his response to the question. A more complete discussion of the differences in rationale between the two concepts of programming is given by Lumsdaine (1960b).

The next paper in the military group, by Briggs, describes two self-instructional devices. The first (Subject-Matter Trainer) has already been briefly discussed (see Figure I, p. 7). An interesting aspect of the conception underlying this device is that some of the variation in amount of stimulus support the student receives at various stages of learning is provided by variations in the mode of operation of the device—rather than by the program content. For example, the amount of prompting in the Subject-Matter Trainer can be varied by changing the mode of operation so that less guidance for the selection of the correct answer is given as training progresses. Several earlier papers on the Subject-Matter Trainer and a study comparing learning under its different modes of operation are abstracted in Appendix I (Briggs and Besnard, 1956; Besnard and others, 1955; and Irion and Briggs, 1957). The other device described in this paper is a four-category multiple-choice device somewhat similar to the early Pressey machines. It incorporates, like Pressey's 1927 device, a provision for dropping mastered items out of the series (also a feature of the constructed-response device described by Skinner, 1958c). A variable-program microfilm device for use in conjunction with an equipment mock-up in teaching troubleshooting for complex equipment has also been described by Briggs (1956a,b).

To extend this work by Briggs and Crowder, a full-scale project on self-instructional techniques was initiated by R. M. Gagné and A. A. Lumsdaine at the AFPTRC Maintenance Laboratory in 1956. This project was aimed at both practical development and supporting research. Cessation of the work begun under this project was forced by the discontinuance of AFPTRC in 1957. Plans for developing a full maintenance training course conducted solely by self-instructional methods had to be abandoned; and further work of this kind was delayed until nearly two years later when the Air Force Cambridge Research Center initiated a project to develop self-instructional training programs for SAGE operators, under contract with the American Institute for Research. Recent in-service studies by Air Force agencies include those of Mayer and Westfield (1958) and Hatch (1959). Currently, several other projects on self-instructional techniques are also being sponsored by federal agencies, including the Air Force (Office of Scientific Research and Wright Air Development Division), the Office of Naval Research, the Human Resources Research Office of the Army, and the United States Office of Education.

The two papers by Rothkopf are products of the work initiated at the AFPTRC Maintenance Laboratory. In the first of these papers Rothkopf presents a schema in terms of which the functional elements of a complete self-instructional system may be analyzed. He also indicates some of the emphases for research to which this conception leads. Another schema for identifying functional elements and interrelationships of a teaching-machine system is presented, from the engineering standpoint, in the paper by Weimer of RCA Laboratories. The other paper by Rothkopf describes two experimental devices constructed for study of variables influencing self-instruction. Rothkopf identifies some of the research problems in the area and makes recommendations concerning instrumentation for research. The 1959 paper by Briggs is concerned with complex kinds of tasks for which self-instructional techniques can be adapted—such as tasks involving both verbal and psychomotor components. Proposed self-instructional devices for such tasks have been described in more detail by Briggs and Duvall (1957).

Systems theory
Servo-mechanism

An independent and unusual approach to teaching-machine concepts is represented in a series of papers by Gordon Pask, two of which are presented here. Pask's approach is different in origin from most of the other work reported in this volume, deriving in part from concepts of automatic control, cybernetics, and self-organizing systems. Several other articles by Pask are abstracted in Appendix I. The first of the two articles included here deals with a specific class of teaching machines designed for teaching perceptual-motor skills of the type involved in operating keyboard machines such as typewriters and card-punching equipment. For such tasks, Pask describes a device (known as SAKI) which varies the sequence of practice materials as well as the timing requirements and prompting features of the task in accordance with the proficiency attained by the trainee. A somewhat simplified commercial version of this device is currently manufactured by the Rheem-Califone Corporation and is illustrated in the figure included in Pask's article.

Teaching-machine systems which are capable of adapting to the responses of the learner in the full sense that the devices "learn" about the idiosyncrasies of the individual subject (and govern the program of materials presented to him accordingly) are envisaged and discussed in terms of a game-theory model by Pask in the second of these papers. Although such devices are outside the main stream of current work on teaching machines, they are of considerable interest as possible future developments. The theory and technology underlying such potential developments are fairly complicated, and a full presentation of them is not considered to be within the scope of this book. Accordingly, Pask has prepared this paper as a relatively nontechnical account of some of the concepts involved, and the editors have included it here to give the reader a picture of this speculative area in the pioneering of advanced teaching-machine technologies. It is interesting to note that both Rothkopf and Pask introduce the idea of "a little man in a black box." Rothkopf (1958) employs this notion as a working procedure in which the experimenter simulates machine functions [cf. Coulson and Silberman, p. 455] in order to develop functional specifications for automated devices. Pask, on the other hand, utilizes the notion as a heuristic for developing a conceptual model of an adaptive teaching machine.

The paper by Ramo stresses the impact of advancing technology, both in setting requirements for more efficient education and in providing possible approaches to the solution of the problems that are thereby posed. Ramo's article, which first appeared in 1957, stimulated a good deal of interest in the possibilities of automation in education, despite its frankly conjectural and discursive nature. A brief overview of Ramo's paper has been presented by Glaser in one of the introductory papers (p. 27-28).

A later, somewhat similar paper was presented by Ramo at the 1959 meetings of the American Psychological Association. An abstract of this paper is included in Appendix I, together with comments by A. W. Melton and B. F. Skinner, who discussed Ramo's paper at the APA meetings.

Finn's paper, the last in this section of the book, is condensed from a more extensive discussion of the history and role of technology as it may affect education. This paper is particularly of interest in terms of its discussion of some administrative and policy implications for education, resulting from increased centralization of instructional control—which are anticipated with the development of mass-produced instructional tools such as filmed or taped educational television programs and programs for teaching machines.

Teaching and Educational Inventions

IBERT MELLAN

AMONG THE most recent of educational problems, and perhaps the most widely discussed, is that classroom procedure known as instruction by mechanical devices. Since the World War active steps have been taken to equip hundreds of classrooms with appropriate apparatus to demonstrate the part vision plays in the pursuance of efficient teaching methods.

It is interesting to find that the attention given to visual instruction goes farther back in educational history than would be otherwise supposed by this only recent attention to its possibilities.

The writer has been fortunate in uncovering a vast wealth of material hidden and filed away in the United States Patent Office which may throw much light on the way of educational study. The following lists of patents are, for the most part, concerned with teaching and educational devices, appliances, apparatus, etc. The earliest record in the United States Patent Office is the contribution of H. Chard, dated February 16, 1809, who gives us a "Mode of Teaching To Read." S. Randall, dated October 1, 1810, and January 11, 1812, offered a patent entitled "Mode of Teaching To Write."

Although not of the earliest contributors to this patent literature, the distinguished name of Dr. Maria Montessori will be found among them. Her patent, No. 1103369 (1914), offered an educational device for properly training the sense of touch, which, in her opinion, is an essential factor in the proper development of the child.

Curiously, following these first two pioneers, patent literature on education and teaching was noticeably infrequent until 1870, after which the more vigorous output of educational inventions seemed to mark a greater interest in teaching and education generally. Each succeeding year saw a continual increase in the

---

This paper is reprinted from the *Journal of Experimental Education*, Vol. 4, March 1936, with minor changes in format.

output so that up to the present time there are between 600 and 700 inventions issued on the subjects of teaching and education.

To keep pace with the never-ending variety of scientific progress, educators must take advantage of every valuable contribution of modern science and industrial and educational invention. I, therefore, believe that an intelligent study of these patents is necessary for all schoolmen. This is a unique opportunity for educators, with their experience and years of training, to develop an abundance of adequate teaching material from this vast amount of literature which contributes to every type of instruction and to every subject in the curriculum. It may be that many of these patents are useless today because they may have been outmoded in the face of constant advancement in education and teaching methods; however, here is the premise of the educator who must decide between the useful and the unfit material.

Copies of these patent specifications may be obtained from the Commissioner of Patents, Washington, D. C. The cost is 10 cents a patent. In ordering, give patent number and title of the invention.

With the aid of the following key and the patent number an approximate time when the patent was issued can be determined.

PATENT NUMBER	DATE WHEN ISSUED
1	1790
32000	1820
60000	1840
110000	1860
230000	1880
443000	1890
532000	1895
667000	1900
817000	1905
980000	1910
1138000	1915
1364000	1920
1568000	1925
1787000	1930
1892000	1932

PATENTS RELATING TO THE TEACHING OF ARITHMETIC

Teaching Arithmetic, 4632 (1846), 149235; Means of Teaching Fractions, 151971; Device for Teaching the Metric System, 176735; Apparatus for Teaching Arithmetic, 196583; Device for Teaching Involution and Evolution, 209385; Device for Teaching Arithmetic, 214822, 215916; Apparatus for Teaching Arithmetic, 234247; Device for Teaching Arithmetic, 262191; Apparatus for Teaching Arithmetic, 264572; Educational Appliance, 296018; Means of Teaching Fractions, 342651; Device for Teaching Fractions, 356167; Device for Teaching Fractions and Percentage, 383300; Apparatus To Facilitate the Teaching of Notations and Numerations, 384959; Device for Teaching Arithmetic, 389415; Apparatus for Teaching Arithmetic, 390824; Chart for Teaching Addition, 416593; Device for Teaching Numbers to Children, 431102; Device for Teaching Arithmetic, 452302; Device for Teaching Combinations of Numbers, 462376; Device for Teaching Arithmetical Calculations, 502184; Apparatus for Teaching Arithmetic, 588371; Device for Teaching Computation, 629891; Educational Device for Teaching Spherics, 812408; Device for Teaching the Fundamental Operation with Numbers, 816204; Device of Facilitating Teaching of Fractions, 841158; Apparatus of Teaching and Learning Arithmetic, 846484, 856068; Device for Teaching Fractional Value, 1043652; Device for Teaching Fractional, 1098330; Device for Teaching Numbers, 1151279; Device for Teaching Arithmetic, 1129890; Appliance for Teaching Arithmetic, 1174689; Device for Teaching Arithmetic, 1211625; Appliance for Teaching Arithmetic, 1218931; Appliance for Teaching Arithmetic, 1405010; Device Used in Teaching Geometry and Trigonometry, 1541179; Device for Teaching Numbers in Combination, Analysis Factors and Multiple, 1594396; Appliance for Teaching Arithmetic, 1662503; Device for Teaching Arithmetic, 1728584; Apparatus for Teaching Arithmetic, 1730418; Device for Teaching Division, 1818566; 380532; 704979; 708568; 1248238; 666999.

CHEMISTRY

Apparatus for Teaching Chemistry, 242821; Appliance for Teaching Chemistry, 480275.

DRAWING

Device for Teaching Drawing, 171268; Cards for Teaching Drawing, 282659; Device for Teaching Drawing, 471442; Apparatus for Teaching Drawing, 651791; Educational Art Text Sheet, 720187; Charts for Teaching the Reading of Drawings, 1049241; Device for Teaching Drawing, 1617207.

GEOGRAPHY

Apparatus for Teaching Geography and Astrography, 2426 (1842); Teaching Geography, 143934; Educational Device for the Illustration of Longitude and Time, 328629; Educational Globe, 418455; Apparatus for Teaching Astronomy and Geography, 501136; Apparatus for Teaching Elements of Geography, 646582; Teaching Geography, 1042455.

HISTORY

Apparatus for Teaching History, 198749; Chart for Teaching Universal History, 1406173.

MUSIC

Sectional Table for Teaching Music, 176471; Device for Teaching Singing, 181312; Device for Teaching Musical Transposition, 183103, 197497; Device for Teaching Music, 203210, 264932, 367156; Indicator for Teaching Music, 417734; Device for Teaching Music, 435131; Apparatus for Teaching Music, 451010; Device for Teaching Music, 510302; Device for Teaching Vocal Music, 528310; Apparatus for Teaching Melodies, 577667; Device for Teaching Time in Music, 585681; Device for Teaching Musical Notation, 618611; Game-Cards for Teaching Music, 621323; Device for Teaching Music, 632137, 657953; Device for Teaching Music and Singing, 736960; Apparatus for Teaching Music, 752836, 762990; Device for Teaching Music, 788063; Device for Teaching the Playing of Stringed Instruments, 830915; Device for Teaching Music, 964593, 979193; Chart for Teaching Music, 1058831; Apparatus for Teaching Music, 1600052; Apparatus for Teaching Sight-Singing, 1685682; Piano Teaching Mechanism, 1732377; Apparatus for Teaching Piano, 539191; 699510; 685447; 692736; 672678; 675723.

PENMANSHIP - WRITING

Mode of Teaching To Write (S. Randall), filed January 11, 1812; filed October 1, 1810. Plummets of Lead in Teaching Writing (T. Weston), filed June 26, 1812. Diagram for Teaching Penmanship, 44778; Copy-Slips for Teaching Penmanship, 161192; Copies for Teaching Penmanship; 172971; Gages for Teaching Penmanship, 177120; Device for Teaching Penmanship, 188984, 217499; Hand-Guide for Use in Teaching Penmanship, 226942; Appliance for Teaching Penmanship, 364249, 414300; Device for Teaching Penmanship, 503796, 507950, 510372, 726898, 735782, 757383, 764970, 783496, 799152, 809712, 862004, 940744, 972273, 1136450, 1184155; Apparatus for Teaching Penmanship, 592029.

PHYSICAL

Apparatus for Teaching the Art of Swimming, 149249; Apparatus for Teaching Swimming, 206892; Apparatus for Teaching the Art of Boxing, 426978; Apparatus for Teaching Diaphragmatic Breathing, 537516; Apparatus for Teaching Swimming, 563578; Apparatus for Teaching Children To Walk, 658464; Apparatus for Teaching Swimming, 946886; Teaching the Golf-Swing, 748532; Mechanical Figure for Teaching Golf, 1703403; Self-Teaching Walking and Dancing, 1815443.

READING

Mode of Teaching To Read (H. Chard), filed February 16, 1809. Apparatus for Teaching Words and Identifying Object, 569846; Means for Teaching Speaking and Reading, 660255; Table of Movable Characters for the Teaching and Practicing of Reading, 822937; Appliance To Be Used in Teaching Reading, 1010782; Means for Teaching Reading, 1224742; Device To Facilitate the Teaching of Sight-Reading, 1263626; Teaching Reading, 1278425; Teaching Reading and the Like, 1584627; Apparatus for Teaching Reading, 1514270, 1706550.

SPELLING

Apparatus for Teaching Spelling, 52758, 189535; Apparatus for Teaching Word-Analysis, 218306; Kindergarten Apparatus of Teaching Spelling, 342737; Kindergarten Apparatus of Teaching Spelling, 366821; Kindergarten Game of Teaching Spelling, 386845; Kindergarten Apparatus of Teaching Spelling, 371815; Apparatus for Teaching Spelling, 342076; Teaching Spelling, 505807; Teaching Spelling for Kindergarten, 792801; Means for Teaching Spelling and the Like, 1099324; Means for Teaching the Alphabet, 1270566; Means for Aiding in Spelling and Phonics, 1326695.

TYPEWRITING AND BOOKKEEPING

Teaching Business Practice, 534723; Machine for Teaching Touch Typewriting, 823362; Device for Teaching Touch Typewriting, 844025; Device for Teaching Typewriting, 1008591; Means for Teaching Shorthand, 1027514; Apparatus for Teaching Bookkeeping, 1415278; 694944; 898114; 665991; 678618.

TELEGRAPHY

Machine for Teaching Telegraphy, 522454; Teaching and Practice of Telegraphy, 736936. Teaching Reading and Sending of Telegraphic Messages, 769422; Instrument for Teaching and Practicing Telegraphy, 773374; Apparatus for Teaching Telegraphy, 789378; Teaching Wireless Telegraphy, 708688, 956936; Teaching Telegraphic Codes, 956936.

MISCELLANEOUS

Educational Block, 341754, 498636, 456708, 577888, 1354910, 1377261, 1477255, 1527051, 1864703; Educational Ball, Des. 64967, June 24, 1924; Educational Board, 1213690, 1392726, 1406592, 1618952; Educational Cards, 1485145, 1506156, 1528690, 1587391, 1571488, 1599568; Educational Chart, 710653, 778928, 874152, 1587928; Educational Construction Set, 1163254; Educational Folder, 1670254; Educational Frame, 718520; Educational Concentrator, 770841; Educational Game, 1533122, 1551895, 1581178, 1767424; Educational Implement, 966473; Educational Model, 639941; Educational Puzzle, 364465, 740451; Educational Novelty, 971865; Educational Material, 736448, 853042, 1656030, 1868823; Educational Record, 1535188; Educational Instrument, 305585; Educational Work and Play, 1627211; Educational System, 656569; Educational Top, 1491123; Educational Test Sheet, 1586628; Educational Test Device, 1490934, 1843183, 1867511; Educational Device for Object Teaching, 1343721; Educational Device for Teaching Spelling, 1619160; Educational Device for Interchangeable Topics, 1729241; Educational Demonstrating Device, 1802331; Figure for Educational Purposes, 630217, 1841369; Apparatus for Teaching Gun Practice, 732596, 751591; Apparatus for Teaching Shooting, 944368, 707729; Means of Teaching Sewing, 337609; Teaching Embroidery, 778162, 1532110; Means for Teaching Parliamentary Law, 946868; Means of Teaching Facial Expressions Which Occur in Speaking, 723716; Teaching the Deaf To Hear and the Mute To Speak, 1676716; Means of Teaching Aviation and Testing Aeroplanes, 1007467; Apparatus for Teaching the Art of Aeroplaning, 1018645; Apparatus for Teaching in Kindergarten, 653838; Apparatus for Teaching Projection, 1327474; Implement for Teaching Fingering, 1666464; Magnetic Indicator for Teaching, 126761; Appliance for Teaching Botany, 496256; Device for Teaching English, 1539397; Time Teaching Apparatus, 825902; Teaching the Reading of Clocks and Dials and the Distinguishing of Color, 1143519; Chart for Teaching Physiology, 298746; Puzzle-Card for Object Teaching, 274799; Object-Teaching Frame, 248659; Chart for Object-Teaching, 197279; Educator, 523338, 1536180, 1561744, 1587864, 1867876; Educational Article, 1502006; Educational Amusement Device, 771394, 973186, 1075248, 1120681, 1386248.

PATENTS INDIVIDUALLY TITLED "EDUCATIONAL APPARATUS"

1370826, 1384801, 1392258, 1394620, 1417434, 1428456, 1433852, 1476671, 1516097, 1523188, 1531070, 1617272, 1634289, 1636234, 1662272, 1700946, 1705315, 1766355, 1810745, 1807615, 1831383, 1839558, 1860895, 1862872.

PATENTS INDIVIDUALLY TITLED "EDUCATIONAL APPLIANCES"*

281770, 309064, 375095, 385046, 388486, 401043, 421044, 466296, 445782, 446468, 448019, 463475, 468475, 472419, 474773, 476468, 480164, 484431, 487695, 496993, 500824, 501675, 505667, 521360, 528010, 530450, 511470, 574815, 577196, 536497, 547217, 553533, 555026, 556431, 563872, 564396, 560964, 572972, 589180, 595455, 600234, 602222, 610277, 617883, 618198, 622551, 623558, 626423, 627365, 629046, 636182, 641151, 641738, 641739, 645440, 646661, 667397, 669878, 672062, 674482, 690446, 690664, 692019, 691113, 706463, 713085, 715224, 723425, 730859, 746305, 747711, 777268, 791743, 791709, 802807, 811169, 814653, 819847, 831154, 832331, 853756, 864090, 873667, 882463, 884863, 885273, 898754, 932852, 951606, 958987, 964064, 967591, 971185, 977793, 997986, 1013856, 1018146, 1028212, 1050327, 1054890, 1084370, 1093690, 1099372, 1158774, 1163125, 1129019, 1138807, 1141340, 1141409, 1150550, 1142651, 1168949, 1182636, 1183068, 1183570, 1189233, 1169510, 1204089, 1204854, 1206795, 1213839, 1228391, 1229142, 1240031, 1243957, 1244000, 1253908, 1256224, 1256997, 1257655, 1262269, 1264449, 1260601, 1269713, 1279504, 1269384, 1270668, 1275955, 1286158, 1286232, 1286631, 1289743, 1289849, 1281295, 1291045, 1294126, 1295404, 1305742, 1310997, 1315478, 1322261, 1329850, 1329896, 1335126, 1343095, 1364778, 1368192, 1369640, 1375308, 1391986, 1435660, 1456395, 1465699, 1470845, 1471437, 1479147, 1480458, 1483916, 1498121, 1502991, 1511124, 1535056, 1538929, 1538930, 1539194, 1543067, 1582810, 1587026, 1603201, 1607329, 1624450, 1630939, 1664842, 1673166, 1679536, 1689422, 1694405, 1696937, 1699289, 1704297, 1742781, 1751106, 1764446, 1677521, 1767774, 1775603, 1784952, 1801724, 1802492, 1811105, 1842165, 1857009, 1869522, 1874279.

---

* These may refer to different subjects of the curriculum.

PATENTS INDIVIDUALLY TITLED "EDUCATIONAL DEVICES"*

196532, 288628, 302194, 335838, 383389, 445759, 544714, 618114, 624614, 625881, 641283, 641683, 649054, 663287, 676313, 677652, 680695, 683267, 683171, 687288, 687570, 688388, 696690, 701997, 711486, 711879, 720510, 731175, 755837, 763926, 774998, 784145, 793676, 793767, 794005, 795855, 801316, 832871, 838840, 871934, 886172, 889515, 892715, 894043, 925716, 927499, 935515, 940093, 947064, 969309, 969429, 1023586, 1028378, 1038332, 1041059, 1043596, 1052460, 1053598, 1071358, 1077515, 1085405, 1087186, 1100362, 1103369, 1113237, 1132409, 1136663, 1139256, 1142947, 1148616, 1163184, 1186267, 1184326, 1170537, 1196099, 1209612, 1218993, 1228197, 1232021, 1233544, 1240556, 1245475, 1305449, 1321292, 1332761, 1346929, 1349930, 1350237, 1354692, 1356929, 1359115, 1383097, 1385096, 1394305, 1396379, 1400887, 1378874, 1405063, 1405193, 1414467, 1417828, 1419882, 1428206, 1426997, 1437037, 1445819, 1455522, 1446941, 1457223, 1457468, 1469919, 1477322, 1479423, 1479876, 1484883, 1485190, 1486690, 1494872, 1497150, 1506210, 1509889, 1519426, 1521491, 1523047, 1530418, 1533437, 1539909, 1535706, 1541795, 1542031, 1549673, 1559665, 1560994, 1561447, 1562518, 1573358, 1578665, 1583061, 1586960, 1587685, 1581390, 1597177, 1598499, 1595115, 1597562, 1605697, 1614390, 1617831, 1617659, 1681560, 1621262, 1634713, 1641773, 1645008, 1647276, 1664390, 1664808, 1674553, 1678621, 1696988, 1729518, 1732815, 1732980, 1734115, 1735456, 1734544, 1745674, 1750799, 1756208, 1762864, 1781047, 1791982, 1798647, 1804813, 1812110, 1820209, 1833793, 1837194, 1838942, 1840507, 1847815, 1854999, 1855823, 1856650, 1860483, 1864022, 1866133, 1882575, 1884476.

[*Some Further Notes on Educational Inventions.* An interesting example of an early teaching device which provides an ingenious display and response arrangement, but would not be considered a true teaching machine because it does not provide immediate feedback concerning correctness of the response, is the device for teaching spelling patented by Haylcon Skinner in 1866 (Patent No. 52758). In more recent devices the arrangements for providing feedback as to the correctness of the response employed vary with respect to rationale and function as well as with respect to details of the mechanical arrangements for achieving these functions. For example, two recent devices patented in the 1940's present instructional items on film and employ multiple-choice questions. The first device (Miesegaes, No. 2394711, February 1946), is a direct-viewing device not requiring a projection lamp, which reveals the correct answer by means of a shutter after the student has attempted to select the correct answer. In a more complex film presentation device patented by Holt (Re. 23030, August 1948), the film is projected, and the learner's choice of an answer category is automatically scored by means of a photoelectric sensing device which determines the correspondence between the selected answer and a coded spot on a portion of the film frame not visible to the student. An interesting "educational toy" in which feedback or correction is given letter by letter to a composed response is the device patented by Composto (No. 2656617, October 1953). Partial obscuring of an extended verbal answer by use of reverse printing is utilized in the device patented by Shulman (No. 1736552, November 1929). The answer can then be read by viewing it in a mirror. One of several recent devices for revealing correct answers against which to check constructed responses to incomplete statements is the educational workbook patented by Studebaker and others in March 1959 (No. 2877567). Other arrangements for first obscuring, then revealing a correct answer are found in the patents by Mathes (No. 1375308, April 1921) and Poritz (No. 2503130, April 1950). A mechanically scored, multiple-choice "teaching machine" employing several response categories and a score summater was patented by Davis in 1954 (2687579). Patent No. 2877568 (March 1959; Besnard and others) covers various features of the Subject-Matter Trainer described on pages 300 to 302.—Eds.

---

* These may refer to different subjects of the curriculum.

The Tab Item: A Technique for The Measurement of Proficiency in Diagnostic Problem-Solving Tasks

ROBERT GLASER, DORA E. DAMRIN, and FLOYD M. GARDNER

RECENTLY, as part of the development of a battery of tests to measure the proficiency of Air Force maintenance personnel, a type of test item has been developed which may prove to have widespread application in the general area of psychological testing. The Tab Item, so called because of its construction, was originally designed to short-cut the more expensive and time-consuming work-sample type of performance measure without sacrificing to a large extent the added validity that such measures usually possess over paper-and-pencil group tests.

Measurement specialists are constantly faced with the practical limitations of individualized performance testing on actual equipment. These limitations multiply as the equipment under consideration becomes increasingly complex, expensive, and inaccessible—as in the case of present-day military instruments. As operational requirements render on-the-job proficiency measures almost impossible, the measurement specialist is forced to invent other techniques by means of which job performance can be closely approximated. In general, these techniques range from the elaborate "mock-up" to the familiar multiple-choice paper-and-pencil test. The first, although usually possessing the advantage of greater validity, is quite costly both in terms of equipment and man-hours of testing time; whereas the second generally achieves economy at the expense of validity. The Tab Item falls somewhere between these two extremes, incorporating

---

This article is reprinted from *Educational and Psychological Measurement*, Vol. 14, No. 2, 1954. This research was supported in part by the United States Air Force under contract AF 33(038)13236, monitored by the Technical Training Research Laboratory of the Air Training Command Human Resources Research Center. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part by or for the United States Government. This paper was originally published as a contractor's report by the Bureau of Research and Service, University of Illinois. Francis G. Cornell was principal investigator. Dr. Damrin is now at the University of Illinois.

SYMPTOM

The screen is dark; no picture appears.

CHECKS

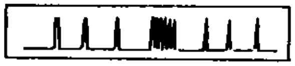
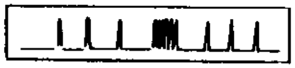
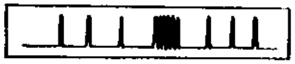
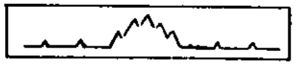
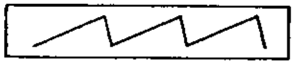
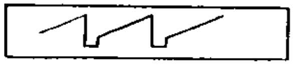
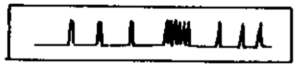
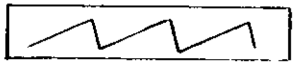
1. TUNING Tune in a station.	Sound appears on all stations but no picture.
2. BRIGHTNESS AND CONTRAST Turn up brightness and contrast controls.	No effect.
3. HOLD CONTROLS Turn up horizontal and vertical hold controls.	No effect.
4. SIZE CONTROLS Turn up horizontal and vertical size controls.	No effect.
5. WAVE FORM CHECKS WITH OSCILLOSCOPE	
A. Detector output	
B. Sound trap output	
C. Video amplifier output	
D. Integrator output	
E. Vertical multivibrator output	
F. Vertical amplifier output	
G. Differentiator output	
H. Horizontal multivibrator output.	

FIG. I. Example of tab item applied to television trouble-shooting

6. OUT OF HORIZONTAL AMPLIFIER Bring a neon bulb on the end of an insulated stick near the top terminal of the secondary of the horizontal output transformer.	Bulb glows brightly.
7. D.C. VOLTAGE MEASUREMENT A. Low-voltage plate supply B. Scope accelerating voltage.	300 volts normal. Zero.
8. SOUND CHANNEL Apply a 4.5 mc. frequency modulated signal to the input of the detector. (Modulation frequency is 400 cycles.)	400-cycle hum appears in loudspeaker. No picture appears.
9. VERTICAL SWEEP Connect headphones across vertical sweep coil.	Loud, rough, roaring sound is heard.
10. HORIZONTAL AMPLIFIER INPUT Connect headphones across the input of the horizontal amplifier.	A very high-pitched tone is heard.

UNITS

1. Detector	NO	7. Picture tube	NO
2. Sound trap	NO	8. Integrator	NO
3. Video amplifier	NO	9. Vertical multivibrator	NO
4. Brightness control	NO	10. Vertical amplifier	NO
5. Low-voltage plate supply	NO	11. Differentiator	NO
6. H.V. rectifier	YES	12. Horizontal Multivibrator	NO
		13. Horizontal amplifier	NO

FIG. I. Example of tab item applied to television trouble-shooting (continued)

certain of the advantages of each while eliminating some of their disadvantages. Whether this device possesses merit is an empirical question which cannot be answered at this time. It is believed, however, that the technique has sufficient logical validity to warrant its preliminary description.

GENERAL DESCRIPTION OF THE TAB ITEM

The Tab Item was originally developed as one of a series of measures of a mechanic's proficiency in maintaining a complex radar-computer system. Specifically, it was designed to measure a mechanic's proficiency in solving problems of a "troubleshooting" type related to the maintenance of this system. Troubleshooting involves the detection and isolation of a defective unit which is causing a particular symptom of set-malfunctioning. Such behavior involves a kind of diagnostic problem solving in which the technician studies the symptom and then performs a series of procedures in which the results of one procedure yield information which supplies a cue for the selection of the next. In this manner the technician obtains the information necessary for him to make a diagnosis. D. G. Ellson, in a personal communication, has called the Tab Item "a test item with feedback."

The troubleshooting of a television set will be used as an example in this description. Although the equipment with which the example is concerned is far less complex than that for which the Tab Item was originally devised, it serves, nevertheless, to illustrate the essential characteristics of the item. The particular illustrative example chosen is restricted to the area of troubleshooting. The wider applicability of the item in other areas will be mentioned later.

A Tab Item consists of several pages (usually three or four) which open so that all of the contents are in view at the same time. As shown in Figure I, these pages contain:

1. A description of a type of set-malfunctioning — the symptom
2. A series of check procedures which might be employed by mechanics to determine the cause of this malfunctioning
3. A list of units or components in the set which might, if defective, cause the described malfunctioning.

The rectangles shown in Figure I to the right of each check procedure and each unit or set component represent perforated tabs which can be pulled out or torn off the page.[1] The information within these rectangles is thus not exposed to the examinee at the time of administration, such information being covered by the tab. Under the tab which accompanies each check is a verbal or diagrammatic description of the results the examinee would obtain if he actually performed the procedure on an operating television set which contained the trouble simulated in the item. For instance, under the tab opposite Check 5A there is a picture of the test oscilloscope screen as it appears after Check 5A has been made; under the tab opposite Check 7A is the voltage reading that would be obtained if Check 7A were made; under the tab for Check 1 is a description of how the sound and television screen would appear if Check 1 were made.

Under the tab to the right of each unit of the set listed is either the word "YES" or the word "NO." It will be noted that only the tab opposite Unit 6 has a "YES" written under it, indicating that there is but one correct answer to this problem.[2] In addition to the material shown in Figure I, a diagram of the system is included with each item, together with any other information to which a repair man would normally have access on the job.

It should now be evident that the Tab Item attempts to reproduce the troubleshooting situation on paper. The test behavior required of the examinee appears to approximate closely the more essential aspects of the behavior required by his job. In the job situation a mechanic generally listens to (or reads) a report of set-malfunctioning. He then proceeds to check out the system in an attempt to discover the cause of the trouble. In the Tab Item the mechanic reads a description of the malfunctioning, and then, rather than actually performing various checking procedures, he pulls the tabs of those checks he would make if he were troubleshooting a real set. Whenever he pulls a tab (i.e., makes a check), he uncovers the information he would obtain if he actually performed that check on a functioning system.

---

[1] For tryout purposes a cardboard square can be stapled to the page to form the tab.

[2] It is possible to insert "double troubles," (i.e., to have two defective components), but this involves additional complexity which will not be discussed in this paper.

In taking a test consisting of one or more Tab Items, the examinee is instructed to locate the defective component by making as few checks as possible. He is free to select any single check, or any combination of checks he thinks necessary to obtain the information which he needs to make a diagnosis, i.e., to isolate the trouble to a specific unit. When he reaches this point, he pulls the tab beside the selected unit and is informed as to the correctness of his diagnosis by a "YES" or a "NO" under the tab. If he finds a "YES," he knows that he has correctly solved the problem presented in that item and proceeds to the next item in the test. If he finds a "NO," he is free to go back and obtain additional information by performing further check procedures before making a second diagnosis. The examinee is specifically instructed that he will be penalized for making erroneous diagnosis, (i.e., pulling "NO" unit tabs). This is done to prevent the location of the "YES" component by trial-and-error selection of the unit tabs.

The words "YES" and "NO" duplicate the on-the-job procedure of checking a component which the mechanic has reason to believe might be faulty. For example, a tube might be tested and found to be dead thus indicating that "YES, this is the cause of the trouble." On the other hand, the tube might test out normally, thus indicating that "No, this is not the cause of the trouble." In such a case the mechanic must continue troubleshooting. In the Tab Item, as on the job, the examinee works until he identifies the defective unit.

CONSTRUCTION OF THE TAB ITEM

As stated previously, the material presented in a Tab Item consists of the following:

1. The symptom of set-malfunctioning
2. The check procedures
3. The results of the check procedures (under tabs)
4. The possibly defective components.

Specific considerations which pertain to the construction of each of these parts will now be discussed.

A. The Symptom of Set-Malfunctioning

The description of the malfunctioning must be clearly and carefully stated. Ambiguity in the description forces the examinee to perform checks for the purpose of clarifying the exact nature of the malfunctioning rather than for the purpose of isolating the defect. If this occurs, a measure of his ability to understand what the item-writer means is obtained in addition to a measure of his ability to troubleshoot. In some situations the deliberate introduction of ambiguity in the description may be desirable, but in the development of this type of item to date this has been kept at a minimum. Wherever possible, a picture or a diagram showing the exact appearance of the trouble should be included. In the example used in this paper the television screen is completely dark, hence a picture is unnecessary. But for a type of malfunctioning in which the image on the screen was blurred and fuzzy, a picture would be presented.

B. The Check Procedures

The checks included in each item must present a wide variety of troubleshooting procedures, from those which are very relevant to the detection of the cause of the trouble involved to those which are quite irrelevant. This serves the double purpose of (a) presenting the examinee with a test situation that retains some of the unstructured aspects of the actual job situation and (b) permitting the wide range of scores necessary for discriminating among several levels of proficiency. These considerations have led to the inclusion of checks of the following types, classified according to the information which they provide:

1. *Relevant.* Information that is pertinent or essential to the solution of the problem, i.e., the detection of the faulty component (A very proficient mechanic would utilize these procedures.)
2. *Additional.* Extra information that a less proficient mechanic might need for solution of the problem but that a highly proficient one could infer from either the description of malfunctioning or the relevant checks
3. *Redundant.* Information which duplicates the information given by another check

4. *Inadequate.* Information which is contained as part of the information given by other more comprehensive checks of the relevant type
5. *Irrelevant.* Information which is completely unrelated to the solution of the problem. (Only mechanics with a very low degree of proficiency would utilize these procedures.)

C. The Results of the Check Procedures

As in the description of malfunctioning, the results given under the tab for each check in the item must be unambiguous and technically accurate. To help insure freedom from ambiguity, a picture or diagram is given wherever possible. The use of a picture or a diagram also serves to lessen the importance of the examinee's ability to visualize verbally presented material. If a result is of a type which an examinee could not be expected to know offhand but would have to look up in a maintenance manual, the obtained result is shown in comparison with the normal result or it is stated that the obtained result is within normal tolerance. (See results of Check 7A.) To insure the technical accuracy of each item, all checks are usually performed on a functioning system into which the defective component given in the item has been introduced.

D. The Possibly Defective Components

Considerations similar to those mentioned for the construction of the check procedures are necessary in selecting the units to be included. Those provided from which the examinee is to select in making his diagnosis are of the following kinds:

1. *Correct.* The component that is causing the described trouble.
2. *Plausible.* Components which could cause the described trouble, but which can be eliminated by performing certain combinations of checks given in the item. For each plausible unit there exists one or more check procedures, which, if performed and if properly interpreted, will permit the examinee to eliminate that component from consideration.
3. *Irrelevant.* Components which could not possibly cause the described trouble.

SCORING THE TAB ITEM

Several methods of scoring the Tab Item have been considered. . . . A proficient maintenance mechanic may be assumed to be one who will detect the source of system-malfunctioning with a minimum of activity. On this assumption, in taking a test of one or more Tab Items, he will select a minimum number of relevant check procedures and proceed immediately to the correct solution. A poor mechanic, on the other hand, may be considered as one who will need more information, will perform nonessential procedures, and will make erroneous diagnoses. Accordingly, the simplest method of scoring the item is a straight frequency count of the number of tabs pulled. In this method, the greater the score the poorer the performance.

A more complex method involves a weighted error count wherein the procedures and units which can be selected are weighted in inverse relation to their relevance in isolating the defective unit. The weights are based upon the kinds of checks and components which the examinee selects. (See B and D under Construction of the Tab Item.) With respect to the check procedures, relevant checks are weighted minimally and irrelevant ones maximally. Additional checks receive a weighting intermediate between these two. Redundant and inadequate checks by definition are meaningful only in terms of related procedures and thus are weighted only as they appear in predetermined combinations. For example, a redundant check receives extra weight when the procedure with which it is redundant has also been performed. With respect to the units, the relevant unit receives a zero weighting and irrelevant units receive maximum weighting. Plausible units are weighted as they occur in combination with the checks that have been selected. If a "NO" unit tab is pulled and the check procedures which relate to it have not been pulled, the unit is weighted maximally on the assumption that the examinee is guessing.

A somewhat different technique of weighting can be based on the seriousness of the error made. In this instance, "seriousness" is defined in terms of those procedures which are unnecessarily time-consuming or are damaging to the system.

Finally, it should be pointed out that in some situations where a Tab Item might be employed, the sequence of the procedures and units selected might be useful scoring information. In such cases a form can be developed on which a record is kept of the order in which an examinee has pulled the tabs.

PROBLEMS OF APPLICATION

One of the most obvious limitations of the Tab Item is that it does not measure the manipulative skills involved in the performance of a procedure. In addition, it makes heavy demands upon the examinee's verbal ability and upon his ability to visualize the functioning of a piece of complex equipment when it is not actually before him. A particular set of behaviors is required for a mechanic actually to perform a series of control settings and then measure a specific voltage output, while a somewhat different set of behaviors is necessary for him to read that certain control settings have been made and then select a particular procedure. This discrepancy between test behavior and job behavior is not unique to the Tab Item, and validation data usually inform us of the seriousness of this discrepancy.

Since a single Tab Item yields a range of scores rather than the usual pass-fail indication, some problems arise concerning the application of conventional item statistics. It is possible to employ the techniques based on dichotomous responses by establishing an arbitrary pass-fail cutting point in the range of item scores, but such dichotomizing of data may sacrifice useful information. An alternative procedure for computing an index of Tab Item difficulty is to use the average number of pulls required by a group of examinees for each item. Subtracting from this average score the minimum number of pulls necessary to solve an item yields the average amount of the deviation from a perfect score. This, or some function of it, can be taken as an index of item difficulty. The minimum number of pulls is determined on the basis of expert opinion and analysis of item responses. For the computation of item validity there appears to be no special problems. The entire range of item scores may be used in correlations with criterion scores if desired.

In conclusion, it seems reasonable that the technique of the Tab Item might be applicable to a variety of subject matters. Its use for the measurement of the proficiency of a maintenance mechanic has been illustrated here. Achievement in laboratory

procedures might be tested in this way. A Tab Item test might be developed, for example, for a course in qualitative analysis in chemistry where the selection and employment of a series of procedures is an important aspect of the course. Proficiency in medical and psychological diagnosis might be tested by adapting this technique to measure the ability of a clinician to perform the procedures necessary for correct diagnosis. In general, the Tab Item is an applicable technique for the measurement of behavior which involves the serial performance of a set of procedures where the performance of one procedure yields information which supplies a cue for the selection of the next and subsequent procedures.

Automatic Tutoring by Intrinsic Programming

NORMAN A. CROWDER

AUTOMATIC tutoring by intrinsic programming is an individually used, instructorless method of teaching which represents an automation of the classical process of individual tutoring. The student is given the material to be learned in small logical units (usually a paragraph, or less, in length) and is tested on each unit immediately. The test result is used automatically to conduct the material that the student sees next. If the student passes the test question, he is automatically given the next unit of information and the next question. If he fails the test question, the preceding unit of information is reviewed, the nature of his error is explained to him, and he is retested. The test questions are multiple-choice questions, and there is a separate set of correctional materials for each wrong answer that is included in the multiple-choice alternative. The technique of using a student's choice of an answer to a multiple-choice question to determine the next material to which he will be exposed has been called "intrinsic programming."

The intrinsic programming technique can be implemented with a number of specific devices. The simplest device, which I will call a "TutorText," is a specially prepared book in which each answer choice is identified with a page number. The reader, choosing a particular answer, turns to the page number given for that answer. There he will find either the next unit of information and the next question, or, if the answer he chose was incorrect, he will find the correctional material appropriate to the answer he chose. He will then be referred to the original choice page to try again. The page numbers in the book are assigned essentially at random, and the reader, therefore, cannot progress from one page to the next except by making an active choice of an answer.

A more sophisticated device for handling intrinsically programmed materials is a microfilm-presentation device, which we call the "Tutor." The Tutor presents the material in microfilm (still) and/or motion picture form and prints on paper tape a complete record of the student's sequence of choices and of the amount of time spent on each. A more detailed discussion of both the Tutor and the TutorText is given below.

There are a number of other possible devices that can be built to handle intrinsically programmed materials. For example, such material can be adapted to an aural presentation as well as a visual presentation, or combinations of both may be used. The basis, or defining feature of the method, however, is the use of the student's choice of an answer to a multiple-choice question to determine the next material to which he will be exposed.

RATIONALE

The person who creates a "teaching machine" might seem to be claiming that he understands human learning in sufficient detail and generality that he can set up, via the machine, those conditions under which efficient learning will inevitably occur. However, we who work with intrinsically programmed devices do not have access to any such educational philosopher's stone. Rather, we suspect that human learning takes place in a variety of ways and that these ways vary with the abilities and present knowledge of different students, with the nature of the subject matter, with a number of interactions between these sources of variation, and with other sources of variability of which we are not even aware. Why, then, do we presume to build "teaching machines" at all?

The answer lies in the feedback control which is inherent in the intrinsic programming method. To predictably achieve a desired result, one must either have an infallible process to bring about the result or one must have a means of determining whether the result has been achieved and of taking appropriate action on the basis of that determination. This latter capability is the basis of the intrinsic programming, or automatic tutor-

---

This presentation of Mr. Crowder's technique of automatic tutoring by "intrinsic programming" was prepared for this volume at the request of the editors, and draws together concepts which have been discussed in several previous papers. Mr. Crowder is currently manager of the Training Systems Department at Western Design, a Division of U. S. Industries, Inc., Goleta, California.

ing, technique. The automatic tutoring devices require the student to respond to the material presented, and the devices, in turn, modify their behavior (exposing new and different material to the student) until the desired result is obtained.

It is worthwhile to point out that the process we are describing is different from that of simply furnishing "knowledge of results" to the student. The test result is used to control the behavior of the teaching machine and need not necessarily be furnished to the student at all. The student may incidentally learn whether a given answer was "right" or not, and presumably this is usually desirable, but the furnishing of this "knowledge of results" to the student is not the primary purpose of the testing. The primary purpose is to determine whether the communication was successful, in order that corrective steps may be taken by the machine if the communication process has failed.

Actually, the paragraph above is unnecessarily restrictive, since the concept of "rightness" or "wrongness" of a student's choice of answers is not always applicable. In general, the intrinsic programming technique allows us to produce a device which "responds appropriately" to what the student does. In the teaching context, "what the student does" is ordinarily appropriately categorized into "makes right response or responses" (since there may be several right responses) or "makes one of several kinds of errors." Therefore we spoke above of proceeding to the next topic or taking corrective action, as would be appropriate in the teaching context. However, the basic technique is applicable to quite a variety of situations in which a device is required to interact with a human being. TutorTexts have been written which play bridge with the reader, and the Tutor could be programmed to play quite a powerful game of chess if we were allowed to restrict the human player's choice of moves to the relatively few sensible ones available at any given time.

A final word on the origin of the term "intrinsic programming," and we will turn to specific discussion of devices and techniques. Evidently any teaching machine that employs a flexible program must have some means to control the exposure of the material on the basis of the behavior of the student. Gustave J. Rath, Nancy S. Anderson, and R. C. Brainerd (1959) have described an experiment in which the student's response was automatically fed into an IBM 650 computer, and this information, along with stored information in the computer, determined the next material to which the student would be exposed. In this case the programming device is an external computer and the method, in our terms, would be one of "extrinsic" rather than "intrinsic" programming. The term "intrinsic programming" simply refers to the fact that the necessary program of alternatives is built into the material itself in such a way that no external programming device is required.

TECHNIQUES

The variety of types of programs that can be prepared in intrinsically programmed form is limited only by the ingenuity of the programmer. The technique itself will accommodate any branching program that depends on the history of the student's choices, although, in practice, practical limitations on the bulk of the material to be prepared must be considered. The simplest form of program step is one in which each wrong answer refers the student (after further discussion) back to the original choice page to try again. An example of such a step [from Page 101 of such a book] would be:

> Page 101
>
> Now, you recall that we had just defined
>
> $$b^0 = 1$$
>
> for any b except where b = 0. We had reached this definition by noting that our division rule,
>
> $$\frac{b^m}{b^n} = b^{(m-n)}$$
>
> will give b^0 as a result if we apply it to the case of dividing a number by itself. Thus,
>
> $$\frac{b^3}{b^3} = b^{(3-3)} = b^0$$
>
> but $\frac{b^3}{b^3}$, or any number (except 0), divided by itself equals 1, so we defined $b^0 = 1$.
>
> We used a division process to find a meaning to attach to the exponent 0. Very well, let's see what other interesting results we can get with this division process. Let's apply our division rule to the case of $\frac{b^2}{b^3}$. What result do we get?

ANSWER	Page
$\frac{b^2}{b^3} = b^1$	94
$\frac{b^2}{b^3} = b^{(-1)}$	115
The rule won't work in this case	119

The student who elects to turn to Page 94 will find:

Page 94

YOUR ANSWER: $\frac{b^2}{b^3} = b^1$

Come, come, now. The rule is $\frac{b^m}{b^n} = b^{(m-n)}$.

Now, in the case of $\frac{b^2}{b^3}$, we have $m = 2$ and $n = 3$, so we are going to get

$$\frac{b^2}{b^3} = b^{(2-3)}.$$

So, 2 - 3 isn't 1, is it? It's - 1.

Return to Page 101, now, and quit fighting the problem.

The student who elects Page 119 will find:

Page 119

YOUR ANSWER: The rule won't work in this case.

Courage! The division rule got us through b^0, where $m = n$, and it will get us through the case where m is smaller than n. In this case we have

$$\frac{b^2}{b^3} = ?$$

and applying the rule

$$\frac{b^m}{b^n} = b^{(m-n)}$$

we get

$$\frac{b^2}{b^3} = b^{(2-3)}.$$

So the exponent of our quotient is (2 - 3) which is -1, isn't it? So just write

$$\frac{b^2}{b^3} = b^{(2-3)} = b^{(-1)}$$

as if you knew what it meant.

Now return to Page 101 and choose the right answer.

And the student who chooses the right answer will find:

Page 115

YOUR ANSWER: $\frac{b^2}{b^3} = b^{(-1)}$

You are correct. Using our rule for division

$$\frac{b^m}{b^n} = b^{(m-n)}$$

in the case of $\frac{b^2}{b^3}$ we get

$$\frac{b^2}{b^3} = b^{(2-3)} = b^{(-1)}.$$

Now, by ordinary arithmetic, we can see that

$$\frac{b^2}{b^3} = \frac{b \times b}{b \times b \times b} = \frac{\cancel{b} \times \cancel{b}}{\cancel{b} \times \cancel{b} \times b} = ?$$

So how shall we define $b^{(-1)}$?

ANSWER	Page
$b^{(-1)} = \frac{0}{b}$	95
$b^{(-1)} = \frac{1}{b}$	104

The step we have just been discussing would be diagramed as in Figure I (p. 292).

Now, let us suppose that we find, on the basis of actual trial with students, that the explanation given on Page 119 is not sufficient. In that case we might rewrite the program to add a subsequence to begin on a new Page 119. We would then write Page 119 as follows:

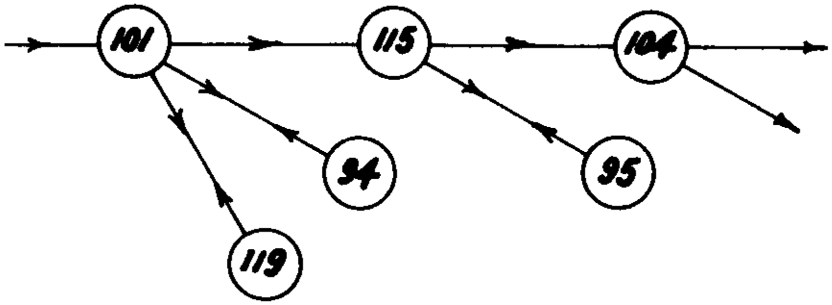


FIG. I. A simple program step

Page 119

YOUR ANSWER: The rule won't work in this case.

The rule is

$$\frac{b^m}{b^n} = b^{(m-n)}$$

And the case we are considering is

$$\frac{b^2}{b^3}.$$

Some of the most interesting cases in mathematics occur when we follow a rule beyond the area in which it was developed. The process is risky but sometimes rewarding. We just reached a definition of b^0 by such a process. So let's forge ahead with the rule in this case and see what happens. In applying the rule, we need only substitute the actual values of the exponents "m" and "n" and do the necessary arithmetic. What do we mean by the letters "m" and "n" in our division rule?

ANSWER	Page
"m" is the larger exponent, "n" is the smaller.	116
"m" is the exponent of the dividend, "n" is the exponent of the divisor.	121

Page 116 might then read:

Page 116

YOUR ANSWER: "m" is the larger exponent, "n" is the smaller. Aha! That was the trouble. Up to now, in our discussion of division, m has always been larger (or equal to) n, so you assumed that this was how we decided which letter referred to which exponent. However, in the statement of the rule

$$\frac{b^m}{b^n} = b^{(m-n)}$$

the exponent m is shown to be the exponent of the *dividend* and the exponent n is shown as the exponent of the *divisor*. Now return to Page 119 and choose the right answer.

Page 121 will read:

Page 121

YOUR ANSWER: "m" is the exponent of the dividend, "n" is the exponent of the divisor.

Very well, in the case of

$$\frac{b^2}{b^3}$$

the exponent of the dividend is 2 and the exponent of the divisor is 3, so our rule

$$\frac{b^m}{b^n} = b^{(m-n)}$$

gives us

$$\frac{b^2}{b^3} = b^{(2-3)}.$$

Now finish the arithmetic and then return to Page 101 and choose the right answer to the question on that page.

Our new structure with the subsequence would be diagramed as shown in Figure II.

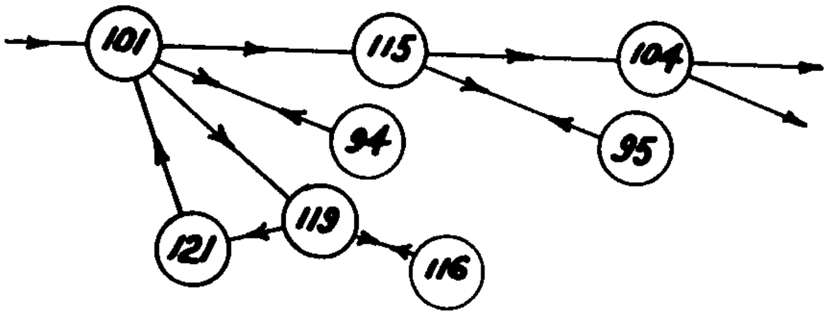


FIG. II. A simple subsequence

Another variation that is frequently useful is one in which a wrong answer returns the student to a previous point in the main sequence to review the preceding development. Such a sequence would be diagramed as shown in Figure III.

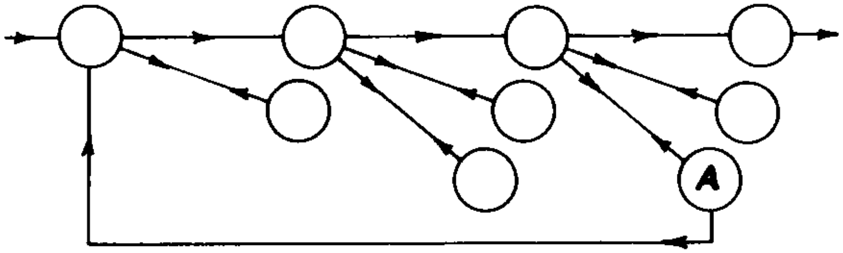


FIG. III. A wash-back program

The student in this sequence who chooses Alternative A in the last step shown is "washed back" two steps in the main sequence.

A variant of the "wash-back" technique is one in which the main sequence of questions is quite difficult and only a few students are expected to get any single answer in the main sequence correct. Those who miss a question in the main sequence are fed to a subsequence which takes smaller steps, as shown in Figure IV:

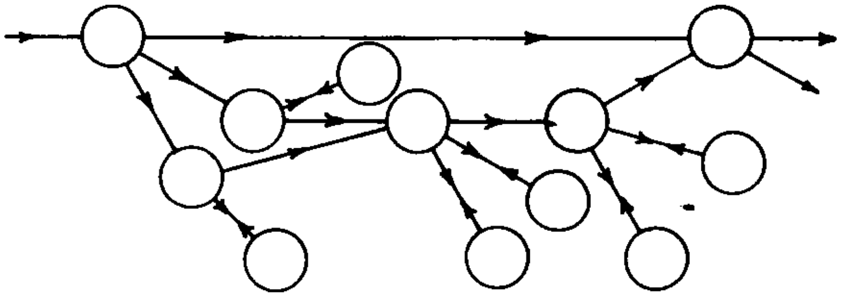


FIG. IV. A wash-ahead program

This type of sequence has been suggested by A. A. Lumsdaine in private communication with the writer. It differs from the subsequence step previously shown, in that the subsequence terminates in the *next* step in the major sequence rather than terminating on the page on which the student left the main sequence.

There is literally no end to the types of sequences that can be developed. A rather fanciful one in which the errors for each question are graded for seriousness, and wash the student back zero, one, or two steps accordingly, is shown in Figure V.

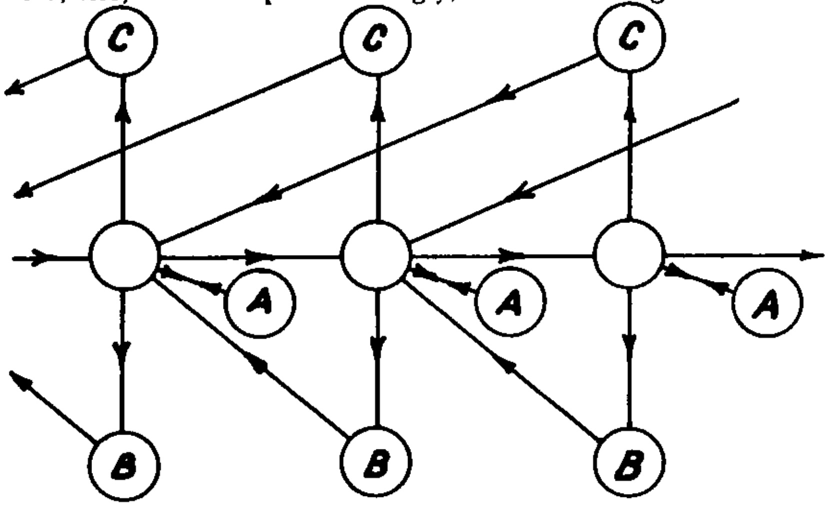


FIG. V. A complex wash-back program

The alternatives marked *A* represent minor errors, from which the student is referred back to the "current" question page; the *B* alternatives wash him back one step, and the *C* alternatives two steps. (The unmarked circles, of course, represent the right answer pages on which new questions are asked.)

A word of caution is required here. It is an interesting exercise to prepare complex forms of sequences and to reason about their properties and probable consequences. The writer who deals with meaningful material will soon discover, however, that such material has a certain stubborn logic of its own and is not always amenable to our predetermined designs. Communication is, at this stage, more of an art than a science, and the conscientious artist will deal respectfully with his subject matter.

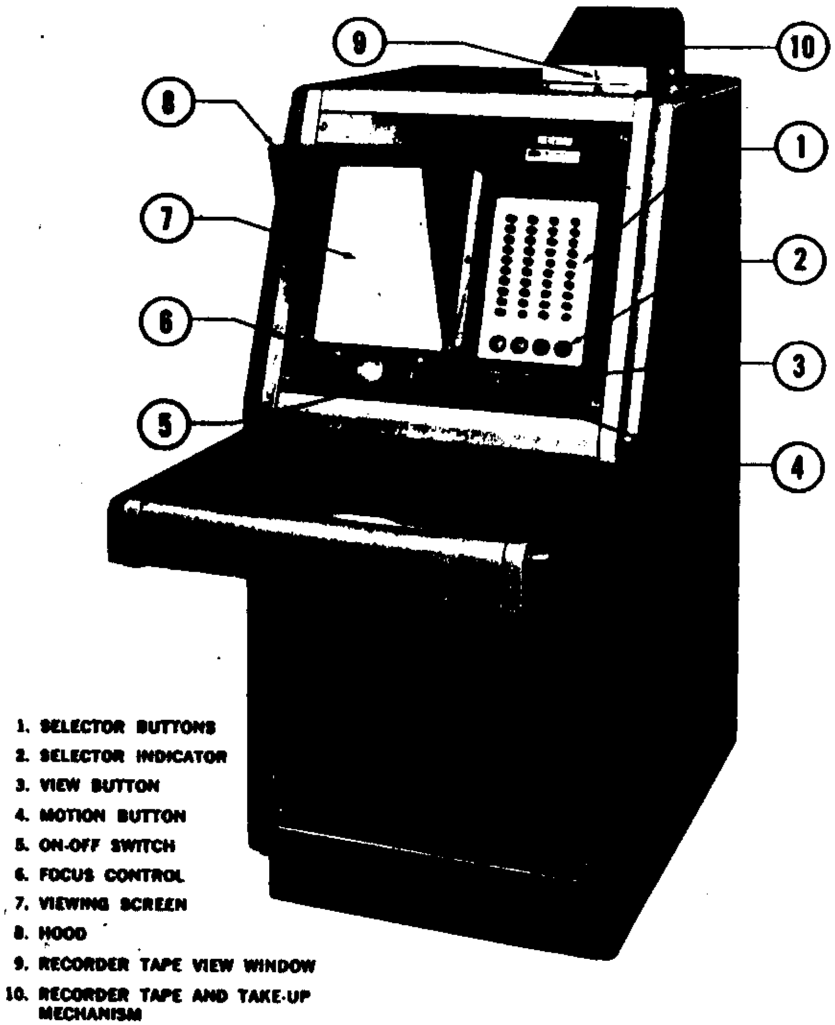
DEVICES

The principal devices used for automatic tutoring at the present time are the Tutor and the TutorText. Sample TutorText pages have been given in the previous section, in connection with programming techniques.

At the outset of the work with TutorTexts, or "scrambled books" as they were originally called, the page numbers were assigned essentially at random, the object being only to prevent the reader from passively scanning the material by not providing a "next page" except as the reader chose an answer to a question. However, as work progressed, it was found that, the order of pages being arbitrary, a number of useful juxtapositions of facing pages could be achieved. The most common example is that of letting a wrong-answer page face the page on which the general discussion of the point in question is given. TutorTexts have been prepared, or are in preparation, in such diverse fields as algebra, trigonometry, physics, chemistry, law, and contract bridge.

The Tutor shown in Figure VI is a complete, recording, automatic tutoring system. The material is handled on 35mm film, and the device has a capacity of 10,000 images, any of which may be called up in any order by entering the appropriate image number in the keyboard. Motion picture sequences may be freely intermixed with still (microfilm) material, but the film moves only when a motion picture sequence is desired. The 10,000-image capacity is therefore adequate for any practical or research purpose.

The student seated at the machine will find the first unit of information on Image 0001, and the question will have choices numbered, say, 0007, 0003, and 0011. Electing one of these answers, he enters that image number into the keyboard, presses the VIEW button, and the required image appears. It may present a new item of information and a new question, or correctional material, just as in the TutorText. When it is desired to show motion, the student will be instructed (by the material in the Tutor, of course) to enter a particular image number into the keyboard, as before, but in this case to use the MOTION button instead of the VIEW button to initiate the search cycle. The motion sequence will then be projected automatically and stop at a question on what the student has seen. Since the film can be stopped at any preselected image and restarted by the student (following directions that appear on the film), any desired degree of control of a film sequence is possible.



1. SELECTOR BUTTONS
2. SELECTOR INDICATOR
3. VIEW BUTTON
4. MOTION BUTTON
5. ON-OFF SWITCH
6. FOCUS CONTROL
7. VIEWING SCREEN
8. HOOD
9. RECORDER TAPE VIEW WINDOW
10. RECORDER TAPE AND TAKE-UP MECHANISM

FIG. VI. The Tutor (Western Design)

The recorder compiles a record of the sequence of images viewed by the student and the amount of time spent on each. In addition, the image number of the last image viewed appears in the recorder view window. This feature is included so that a student who has inadvertently entered a completely wrong number (i.e., a number not included in the choices provided) and thereby "gets lost" can return to the image he was previously viewing.

The advantages of the Tutor derive from the recording capability, the use of film, and the large information storage capacity. The complete recording allows the material to be developed and tested and permits an unprecedented degree of administrative control of the teaching situation. The use of film permits picture sequences to be used and allows accurate color renditions to be economically achieved. This latter point is of some importance in such areas as the training of medical technicians. The use of film is, of course, responsible for the 10,000-image storage capacity which allows us to use any reasonable branching program and still have plenty of space left for motion sequences.

SUMMARY

To sum up what has been said, we approach the design of a teaching machine as a problem in communication. The conditions of the problem are such that the greatest flexibility, both within and between program steps, is required. The within-steps flexibility is required because we wish to communicate complex information to a complex organism, that is, an intelligent human being. To insure the necessary flexibility within steps, the devices provide that any unit of information up to the amount that can be presented on a single page or a page-size viewing screen can be presented at a single presentation. The requirement for flexibility between steps arises because communication may fail, particularly if we are attempting to get the student to flex his mental muscles a little and give him fairly stiff questions. The required flexibility is achieved by the use of the intrinsic programming technique, in either the TutorText or the Tutor.

Two Self-Instructional Devices

LESLIE J. BRIGGS

TODAY THE armed forces employ various types of training aids, devices, and simulators, to promote rapid, efficient, individualized learning of factual material as well as complex intellectual and motor skills. Often these devices provide realistic and effective training in a simulated context to avoid full dependence upon costly, and sometimes hazardous, field training and war games. This paper deals with two special devices intended primarily for promoting learning of technical information.

Self-instructional devices of the kind described in this paper are not in widespread use in public education. This is surprising, especially since much of the material to be learned lends itself readily to written practice exercises adaptable to automatic instruction. Use of such devices would free the instructor from teaching drill material and would permit him to concentrate upon discussion of issues and other activities of broad significance beyond the factual material itself. Pressey (1950) and his students and, more recently, Skinner (1958c) have shown that simple machines can foster effective learning and that they can employ both recognition (multiple choice) and reconstruction (free response) methods of practice. Although there are many issues concerning learning theory and the nature of specific learning materials which legitimately have a bearing on the choice of a multiple-choice versus a free-response device, without further ado it is stated that multiple-choice devices are of great practical utility in many real-life learning situations.

The purpose of this paper is to summarize the features of two

---

This article is reprinted from *Psychological Reports*, Vol. 4, 1958. The devices described were developed while the author was an employee of the Maintenance Laboratory of the Air Force Personnel and Training Research Center, Air Research and Development Command, Lowry Air Force Base, Colorado. (Stipulations relating to proprietary rights concerning these devices are given in a footnote to the original article.) The opinions expressed in this report are the author's; they do not necessarily reflect the opinions of the U. S. Air Force. The devices described have been employed in several research studies (see Briggs and Besnard, 1956, and Irion and Briggs, 1957). Dr. Briggs is now at the American Institute for Research.

self-instructional devices which were developed initially for military training and research purposes but which also could be used for educational purposes. Although the devices described are more complex than many devices developed in the past, they also are more versatile, and to some degree they introduce the potentially useful idea of programming the materials selectively in accordance with the subject's prior performance.

THE SUBJECT-MATTER TRAINER[1]

The Subject-Matter Trainer [see Figure I of paper by Lumsdaine, 1959d; p. 7 of this volume] presents stimulus items one at a time in a small window on the left; *S* may select his answer from among the 20 response items continuously present to view on the right-hand side. Adjacent to each of the 20 response alternatives is a response button to be pressed when *S* has chosen that item as his answer. If he has chosen correctly, a green light next to the answer glows; if he has chosen a wrong answer, a buzzer sounds. Series of 20 stimulus items are presented repeatedly until mastery is achieved. The circuits for this device have been published (Besnard, Briggs, Mursch and Walker, 1955).

The device was designed to permit practice under several conditions or modes of operation, in case there turn out to be differences in relative effectiveness of the modes among *Ss*, stages of mastery, and types of materials. Selection of mode is made by turning a single control knob in the left vertical panel.

The coaching mode—*S*, having observed the stimulus item, presses a special switch; at once the green light next to the correct answer glows, thus showing the student the right answer. *S* then presses an activator control to bring the next stimulus item into view.

The single-error-permitted mode—*S*, having observed the stimulus item, presses the response button adjacent to his chosen answer. If his response is correct, the green light glows, and *S* may go on to the next stimulus item. If the answer is wrong, the buzzer sounds, but immediately thereafter the green light next to the correct answer glows. Thus only one overt error may be made before prompting is given, and *S* must acknowledge the correction by pressing the switch next to the correct answer indicated before he can proceed.

The practice mode—the buzzer sounds after each error made; *S* may make any number of errors on each item. When he makes the right response, the green light glows, and the next item may then be attempted.

The single-try mode—*S* is given only one attempt at each item before he must go on to the next. He observes the green light or the buzzer, whichever is appropriate, after each response.

The paced-practice mode—an electrical timer may be connected to the device to control exposure time of the stimulus item under any of the previously defined modes of operation.

The test mode—*S* is given one try at each item, with no buzzer or light used. Counters tally the numbers of right and wrong responses.

Any question and answer material which can be presented in written, pictorial, or symbolic form may be inserted quickly by changing the cardboard stimulus wheel which revolves under the stimulus window and by placing a new answer panel on the right top surface of the device. Reference materials such as circuit diagrams may be placed in the lid of the device. Such reference materials may show numbered points on a circuit, map, or other data presented, in which case the same numbers are placed on the response panel as the alternatives offered for each question.

Notches are made along one edge of the answer panel in order that the relationship between serial order of the stimulus item and relative location of the right answers may be "scrambled" from problem to problem, or if desired, from trial to trial for a single problem series.

A control knob in the recessed panel on the left vertical side of the device may be turned to employ a red light rather than the buzzer as the signal for a wrong response. The same control knob also permits use of the buzzer for right rather than wrong responses.

It may be noted that the device can be used for learning of paired-associates, with the capability of changing either the serial order of the stimulus items or the relative location of the right responses, or both, by changing the stimulus disk or the response panel. Also, serial or sequence problems, such as steps in a fixed procedure, may be employed. Problems requiring logic, problem solving, or interpretation of data may be introduced by use of supplementary data placed in the lid.

---

[1] This device was developed jointly (Besnard and others, 1955). The model shown in Figure I was fabricated by the Hathaway Instrument Company, under Contract No. AF 18(600)-1058. Patent No. 2877568.

The lower portion of the trainer is simply a storage space for problem materials mounted on the stimulus wheels and on the response panels or trays.

The Subject-Matter Trainer has been employed in two Air Force studies. In one experiment, use of the device by an experimental group resulted in maintenance criterion scores which were significantly superior to scores earned by a regularly trained group (Briggs and Besnard, 1956). Another study revealed the coaching mode to be a surprisingly effective condition for learning of three types of materials (Irion and Briggs, 1957). The single-error-permitted mode ranked second in effectiveness. Use of these two modes early in the learning period followed by the practice mode also proved to be effective. The device appears to offer useful features for conducting research on numerous variables relating to verbal and symbolic learning and retention.

THE CARD-SORT DEVICE[2]

The Card-Sort Device shown in Figure I provides practice modes almost identical to those of the Subject-Matter Trainer. However, each item appears in the window on a separate card, which is later sorted into either one of two retaining wells, depending on the correctness or incorrectness of the first response made to the item. Each item may have from two to four possible answers. Thus the device continuously discriminates right from wrong answers; and on each succeeding trial, *S* is presented only items previously answered incorrectly until there are no cards left in the "wrong" deck. Red and green lights above each answer key furnish the knowledge of results after each response. When an error is made, the red light remains on above the response key, warning *S* not to repeat the error. The red light goes out as the next card is presented. At the end of an exercise, *S* may be given a test trial over the complete deck of cards, or new exercises may be assigned. A maximum of 100 cards may be inserted for each exercise.

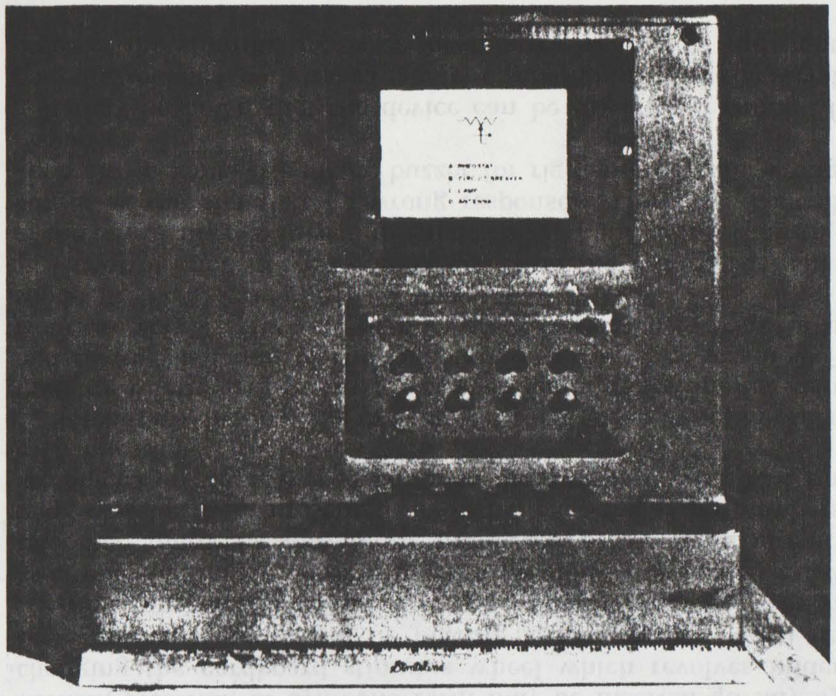


FIG. I. The card-sort device

SUMMARY AND DISCUSSION

Two multiple-choice self-instructional devices are described. These enable *S* to learn correct responses to stimulus items without the aid of an instructor. *S* may work at his own speed, with knowledge of results after making each response. One device, on each practice trial, presents only items previously responded to incorrectly. Multiple modes of operation permit *S* to be shown the correct answers at the outset of practice but require overt responses of him later. The modes may be selected to suit the difficulty of the material and the stage of learning simply by turning a knob. Control over the probability of error may be achieved by selection of the mode of operation to be used and by presentation of only a few alternative answers for each item.

---

[2] This device was developed by the author. The model shown in Figure I was fabricated by Aircraft Armaments, Inc., under Contract No. AF 41(657)-168. Patent pending.

Such devices would relieve teachers of much of the burden of teaching factual material, thus leaving more time for other instructional functions. The devices would permit individualization of instruction, by which bright students could learn more than their classmates, once the required materials were mastered. Sets of learning materials in science and other topics could be provided for those who quickly master the assigned lessons. Since it is not necessary that the teacher be familiar with all the material, a vastly enlarged and enriched learning area is available for the capable student.

The student must respond to each instructional item. No response is ignored, and knowledge of results is given immediately. For all students, bright or dull, the machines are infinitely patient and accurate. They never scold the student, and they never employ sarcasm. They not only warn the student of error, but require him to persist in his active study effort until a correct response is learned to each item presented. Review and scoring are convenient, and people like to work the devices. Finally, while it is practically impossible to "beat the machine" in a test situation, in a training session the machines give the student the amount of help and prompting which is appropriate to his stage of mastery of the material.

Heuristic Discussions of Psychological Bases for the Conduct of Training by Automatic Devices: A Functional Diagram And an Approach to Research

ERNST Z. ROTHKOPF

A FUNCTIONAL DIAGRAM

ONE OF the major objectives of this paper is to propose a functional division among various components of a general automatic training device. This division is not being peddled as a contribution to knowledge in this area. Instead, this division is intended to organize subsequent discussions of research issues arising in the design and use of automatic training devices.

The functional components of a general automatic training device are shown in Figure I. The various components of this schema do *not* represent physical components but, rather, delineate functions which one or more components of a device may perform.

In Figure I, problem source or storage (a) simply represents the electrical or mechanical features of the training device which supply the input for the stimulus presentation gate (f). A problem source may range in its actual form from a box of vocabulary flashcards, or the paper loop of a memory drum, to magnetic memory units containing the problems in analogue form. The problem selector (b) represents that function of the automatic training device which determines the *next* problem (stimulus) event which will be presented to *S* [the subject or learner]. In its actual form a problem selector function may be the serial order of questions on a question wheel as in the SMT [Subject Matter Trainer: see Briggs, 1958], or the page directions in the Crowder automatic tutor books [cf. Crowder, 1960b].

This paper was originally prepared as a staff memorandum for limited circulation at the Maintenance Laboratory of the former United States Air Force Personnel and Training Research Center, April 1957. Dr. Rothkopf is currently at the Bell Telephone Laboratories, Murray Hill, New Jersey.

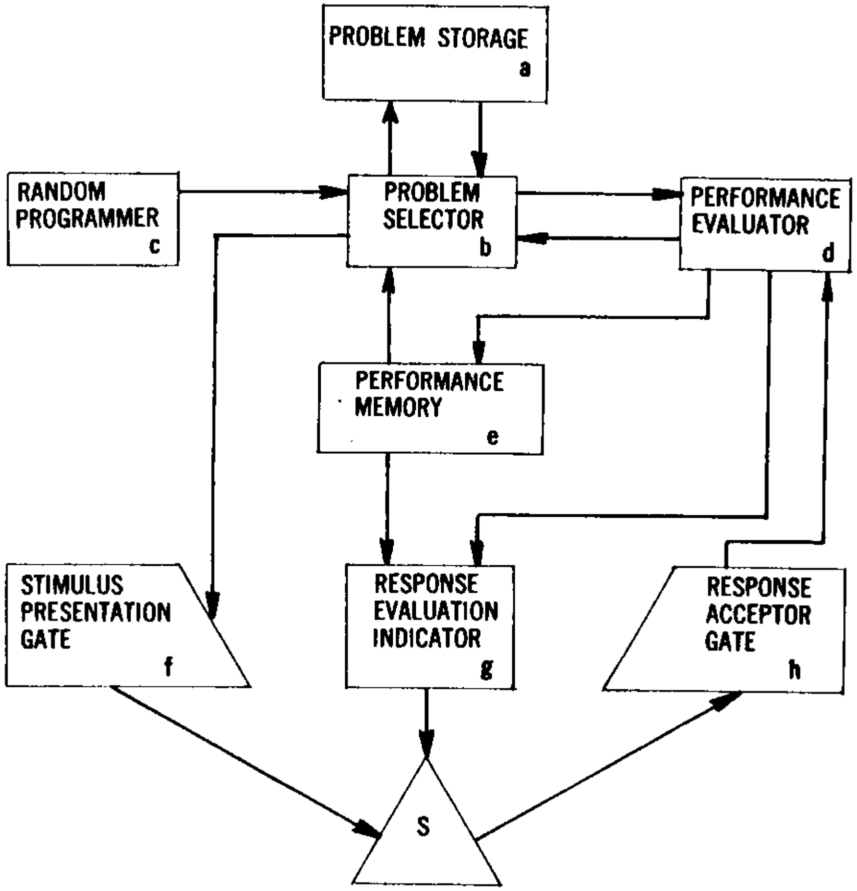


FIG. I. Diagram of the functional components of a general device for automated training. The arrows indicate the flow of signals.

The operation of the problem selector may be determined by three separate functions. These are the random programmer (c), the performance evaluator (d), and the performance memory (e). The need for the random programmer stems from the fact that human subjects tend to associate events in time; i.e., they learn time sequences or serial connections. Such serial habits often interfere with the growth of the habit(s) which are under training. The random programmer is intended to prevent such interfering serial associations by imposing randomness on the problem sequence. Actual examples of a random-programming function are changing question wheels between SMT trials or shuffling unlearned vocabulary flashcards.

The performance evaluator (d) determines whether *S*'s response(s) were appropriate to the problem stimulus. The performance evaluator then supplies a signal for the problem selector. This signal, in turn, wholly or partially determines what problem is to be presented next. In one of Skinner's arithmetic machines (multiple-slide device: see Skinner, 1954) this function of the performance evaluator is carried out by making advance to the next problem contingent on an appropriate response. The same holds true for the SMT in the practice mode. The performance evaluator in some devices also provides signals for S through the response evaluation indicator (g). In the SMT, for example, this signal may be a green light which goes on if S has made a correct response.

The performance memory (e) retains information about *S*'s past performance. The performance memory receives an input from the performance evaluator. The output of the performance memory governs to various degrees the actions of the problem selector. A performance memory function is incorporated in the Shephard-Lewis (1950) paired-associate device when it is used with the "adjusted-learning method."[1] The elimination of items which have evoked a correct response from the presentation sequence is equivalent to a short performance memory function. Similarly, one of Skinner's machines eliminates mastered problems from the practice order.

The stimulus presentation gate (f) converts events, features, or energies of the automatic training device into stimuli to which S can respond. An example of a stimulus presentation gate is the ground-glass screen of the automatic microfilm projector and the attendant projection system. The response acceptor gate (h) offers means by which *Ss* can respond to problem stimuli. In its actual form, the response acceptor gate may range from the coded buttons of the SMT to the constructive response slides of the Skinner (1954) and the Zeaman [2] arithmetic teaching devices. The stimulus presentation function as well as the response acceptor gate are simply the points where the language of the machine is changed into S's language and vice versa.

---

[1] The method of adjusted learning appears to have been first described by Woodworth (1914). The method eliminates a given paired-associate stimulus from practice as soon as some arbitary performance criterion has been reached for that stimulus.

AN APPROACH TO RESEARCH

The preceding pages have been concerned with a brief discussion of various functions which can be occupied by components of automatic training devices. Each of these functions suggests dimensions in which devices may differ from each other. It is the purpose of this section to discuss a general methodology by which empirical solutions to some device design problems may be obtained.

The methodology of research which is proposed here is based on the assumption that it is wasteful and inefficient to design training hardware without specific notions about what the hardware's capabilities must be. In order to determine the desired capabilities of a device or device component, it is proposed that the developmental path of automatic training apparatus pass through the equivalent of the mock-up stage in cockpit design. In this mock-up stage an effort is made to arrive at a constellation of function values which yields good training results. This is done by "simulating" the various functions of a planned training device through use of common laboratory apparatus or other improvisations and by comparing results for the various function values.

An illustration may serve to clarify the general nature of the mock-up technique. Suppose that we wish to develop a paired-associate training device and wish to use a version of the adjusted-learning method (Woodworth, 1914). What sorts of performance memory function should the device have? It would not be difficult to simulate the performance memory function and other features of the machine either using an Osgood screw-type memory drum or by having *E* [the experimenter] sit behind a screen and show stimulus cards to S according to some carefully planned schema. In this manner, the criteria used to eliminate problem (stimulus) cards could be systematically varied. This method could also be used to study the relationship between the criterion used for the elimination of stimulus cards from the practice sequence, and attributes of the training stimuli such as similarity, complexity, etc. The yield of this procedure may include information as to whether the desired device should have a long or a short or an adjustable memory, or what the capacity of this performance memory should be in other respects.

It is the belief of the writer that this general procedure can be applied in at least two ways. In the first, the characteristics of an *entire device* are approached through a series of successive approximations. Each approximation is brought about by some change in the mock-up. Under these circumstances it is likely that the naturalistic methods of the ethologist would be more useful to the research worker than conventional statistical inference.

The second way in which the mock-up procedures can be used in the development of automatic training devices is in maximizing the training effectiveness of the various *individual* functions. It is reasonable to suppose that there exist values for each of these functions which are optimal for particular training materials or purposes. In the best possible state of affairs these optimal values would be deduced from theoretical laws. By use of the mock-up technique in careful parametric studies, the empirical bases for such laws can be provided. Regardless of whether the method is used in the narrow developmental sense or is exploited through extensive parametric studies, the ultimate products of the mock-up procedure are, of course, *functional design specifications* which can be included in development contracts for training-device hardware.

Extensive experience with the technique may also lead to the development and perfection of standard apparatus for "simulating" various device features. This standard apparatus may lead to modular units, each of which is designed to fulfill some specific device function. Such modular units would be aimed at becoming building blocks from which complete training devices for various purposes may be assembled.

The usefulness of the mock-up technique may not be restricted to the development of *automatic* training devices. Probably most part-task trainers and other similar training equipment might also benefit from passing through a mock-up stage during the course of their development. It is not necessary to have a completed piece of troubleshooting training equipment before finding out whether a normal non-normal return from a dial is as effective in practice as many-valued dial indicators. In this fashion, empirical solutions for many design problems can probably be supplied during the period when hardware features of the device under development are being considered.

---

[2] Personal communication to the author.

SUMMARY

This paper discussed nine functions which can be assumed by various components of automatic training devices. These are: (a) problem storage, (b) problem selector, (c) random programmer, (d) performance evaluator, (e) performance memory, (f) stimulus presentation gate, (g) response evaluation indicator, and (h) response acceptor gate. The various functions are suggested as a possible way to organize research problems pertinent to the development of automatic training equipment.

A method for determining optimal values for such functions is proposed. This method, which resembles the mock-up stage of hardware design, simulates various device functions by means of common laboratory equipment and other improvisations.

A Proposed "Automatic" Teaching Device

PAUL K. WEIMER

INCREASING mechanization in teaching is evidenced by the many references to the classroom use of electronic equipment such as television displays, tape recorders, and test-scoring machines. These devices extend the skill of the teacher just as power tools extend the capabilities of a machinist, but there is still much routine in the teaching process even with these aids.

For several years it has appeared to the writer that electronics could make its most significant contribution to education in the form of an "automatic" teaching device capable of carrying out many of the more routine teaching activities and freeing the teacher for matters requiring greater skill. An automatic teaching device, as defined here, requires both the presentation of information and the testing of the student in a controlled feedback relationship. These requirements are most easily met by electronic means. The rate and content of new material to be presented would be determined automatically by the response of the student to frequent tests. Although no two students need follow the same procedure in learning a lesson, the objectives and pedagogical approach could be guided by a "lesson plan" which is supplied to the device in the form of programming information. Each student could progress at the rate of which he is capable without requiring a human teacher to maintain constant control over each step of the process. The writer believes that such machines could be widely used in the schools of the future and would lead to vastly improved instruction at lower cost.

The general concept of teaching devices with feedback has already been utilized by the experimental psychologists in studies

---

Reprinted from the *IRE Transactions on Education*, Vol. E-1, Number 2, June 1958. Mr. Weimer is at the RCA Laboratories, Princeton, New Jersey. In a note of acknowledgment in the original article, the author mentioned that this paper was based on a company memorandum written several years earlier and its being submitted for publication was largely owing to the interest and encouragement of E. W. Herold.

on the science of learning (Skinner, 1954). Devices which permit immediate "reinforcement" or confirmation of a correct response have been shown to promote rapid learning of complex behavior patterns in animals. Skinner (1954) has also proposed simple display-type devices based on this principle for the teaching of children. The use of devices incorporating feedback have not, in the opinion of the writer, received adequate consideration either by the educators whose efforts would be aided by their use or by the electronics engineers who could bring the most modern techniques of information-handling and communications to bear on their design.[1] An exciting opportunity exists for a cooperative effort between educators and the electronics industry which could be of untold benefit to society and, at the same time, might open a vast new market for electronic equipment.

Figure I shows in block diagram the functional relationship of the student with the basic elements of the proposed machine. The block labeled "automatic control center" includes all circuits necessary to select automatically the proper material for presentation to the student in accordance with his test responses and the instructions in the lesson plan. The control center should also provide a timing mechanism for monitoring various steps in the lesson.

The "lesson plan" block in Figure I represents the total information required by the machine to conduct the lesson recorded on such media as magnetic tape, film cards, connector boards, etc. This information includes lesson material, work assignments, and test questions for the student as well as programming instructions for the machine. Many different subjects at a wide range of levels can be taught by a single machine by choice of the proper lesson plan.

The blocks labeled "presentation" and "test" in Figure I represent the transducers used to establish a two-way communication channel with the student. Earphones or a visual display such as a small TV screen would be effective for presentation, but many other possibilities exist. Much of the information presented to the student by the machine will consist of assignment and study aids to assist him in the use of textbooks, maps, laboratory experiments, etc. Questions requiring frequent response by the student will supply to the machine the feedback information necessary for most effective guidance.

---

[1] Since the preparation of this paper a notable exception to this statement has appeared in the form of a recent paper by S. Ramo, "A New Technique of Education." [This was a 1958 reprinting of Mr. Ramo's 1957 paper: see Appendix I.]

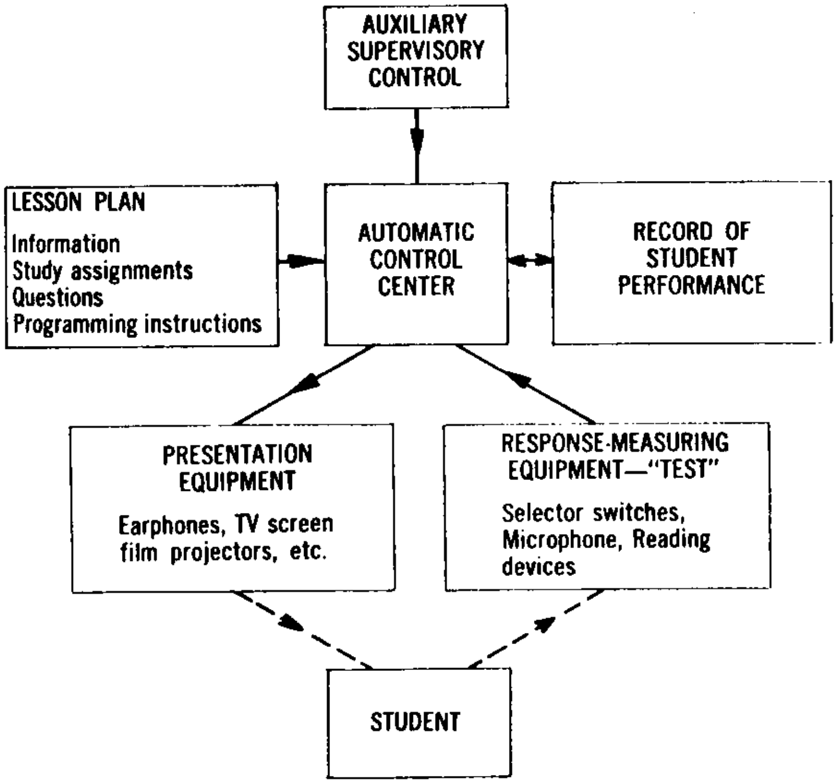

AUXILIARY SUPERVISORY CONTROL
LESSON PLAN
Information
Study assignments
Questions
Programming instructions
AUTOMATIC CONTROL CENTER
RECORD OF STUDENT PERFORMANCE
PRESENTATION EQUIPMENT
Earphones, TV screen film projectors, etc.
RESPONSE-MEASURING EQUIPMENT—"TEST"
Selector switches, Microphone, Reading devices
STUDENT

FIG. I. Block diagram of an "automatic teaching device" for use by an individual student.

The testing function can be accomplished rather simply and effectively by providing a multiple-choice selector switch which each student can manipulate. Other more sophisticated test procedures in which the student responds verbally or in writing can also be visualized. In each case the machine evaluates the student's answer by simply comparing it with the possible answers anticipated on the lesson plan, taking into account, when desirable, other factors such as time for completion and previous errors. The test procedure appears to be quite straightforward when multiple-choice questions are used for testing. By selecting the answer he believes to be most nearly correct, the student simultaneously determines his next assignment and activates the machine to respond to his choice immediately in terms of praise or specific suggestions.

The most effective lesson plans for each subject would evolve from an analysis by skilled educators of the record of the students' performance recorded by the machines. In general, the material for each lesson would be arranged in a logical sequence of increasing difficulty with the proper amount of repetition, review, and drill for effective teaching. At the same time, each lesson plan would include much additional material, with many assignments planned specifically for the use of the superior or the slower students. No one student would be expected to cover all the material included in one lesson plan. The frequent use of perceptive questions would insure that each student receives that portion of the material which would be of greatest benefit to him. The test questions would thus assume a more important role than conventional examination questions, since they are an integral part of each step in the teaching process.

The block labeled "supervisory control" in Figure I indicates methods by which the supervising teacher may supply additional instructions to the automatic control center. By enabling the machine to take into account other factors not already included in the lesson plan, such as personal characteristics or prior experience of a particular student, a high degree of flexibility can be achieved.

While teaching devices incorporating all the features described above would be too complex for general use at present, the writer believes that highly effective results could be obtained with a relatively simple device to be attached to a student's desk. This machine would use earphones for presentation of verbal instructions, multiple-choice switches for test response, and magnetic tape for storing the lesson plan and recording student performance. Indicator lights for registering rate of progress during the lesson could be styled to provide interest and motivation in the young student. A built-in timing mechanism controlled by the lesson plan would develop speed and efficiency in study habits. A wide variety of subjects ranging from elementary to college level could be taught on this machine. The constant challenge and stimulation afforded by cleverly designed lesson plans should make possible the successful application of the device at any age level.

Although the automatic teaching device would appear to offer its greatest advantage in individual instruction, the method

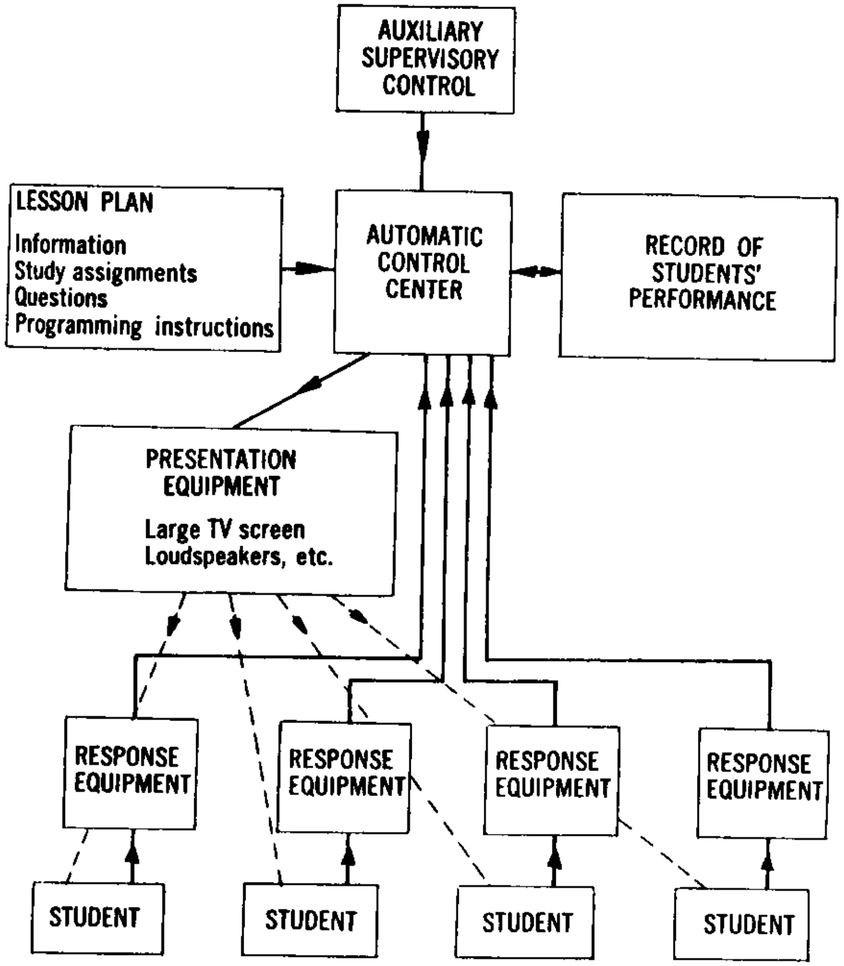


FIG. II. Block diagram of an "automatic teaching device" for use by a group of students.

could also be extended to class instruction, as shown in Figure II. The same presentation is given to all students, but the rate and form of presentation would be determined automatically by the combined response of the entire class. This system might be desirable for certain types of instruction and would permit a simplification of presentation equipment.

The mature development of automatic teaching devices, which might be possible within the next five or 10 years, would offer, in my opinion, the following benefits:

1. Incorporated into the public-school system, they could provide more effective teaching than is now practiced in the average school. Specifically:
    a. They would permit individual instruction for students of all ranges of ability. With present-day classes of 30 or more, no teacher is able to guide the thinking of each student.
    b. They would insure the application of the best known teaching methods for each aspect of the lesson to each child's situation. Even a good teacher cannot be expected to be aware of, or have time to apply, all the details of teaching techniques which have been shown to be best for each lesson.
    c. They would permit a more scientific approach to the improvement of teaching methods by providing more quantitative data on methods and results.
    d. They would free the skilled teacher of much routine activity, permitting more time for inspirational contacts with the student.
2. Automatic teaching devices might provide a means for reducing the cost of education. The large market would result in mass production techniques by competitive suppliers. This factor, combined with the inevitable simplification resulting from technological advances, would suggest that the cost per machine would be considerably less than $1000. The total initial cost of equipping a school of 400 pupils with 100 machines would be less than $100,000.
3. By providing an effective means for obtaining formal training at home, automatic teaching devices should find wide private usage. This could range from occasional homework for backward students to complete courses for children or adults using accredited lesson plans. The machines would provide tabulated evidence of progress to determine course credits.

Automatic teaching devices can be built today which, in the writer's opinion, would function satisfactorily with minor supervision for many subjects ranging from kindergarten to college level. Although their effectiveness would be expected to increase as experience is gained, there are a few subjects especially dependent upon a personal exchange of ideas or skills in which the machines would be at a disadvantage. Their greatest weakness at present would lie not in the presentation of information, but in the forming of an accurate evaluation of the student's understanding. Recent advances in the design of electronic computers suggest good possibilities for improvement of this facility.

The development of the most effective lesson plans would require extensive tests with cooperative effort by educators and electrical engineers. These would be time-consuming and may require expensive experimental models. The rapid obsolescence and maintenance of equipment in the early stages of development might be a problem.

In spite of these limitations, the writer believes the wide use of automatic teaching devices for routine instruction to be inevitable, at first as a teaching aid under supervision in a few subjects and eventually as the central framework for organized instruction in many subjects. To properly carry out such a development, however, will require an extended cooperative effort of electrical engineers and professional educators.

Some Research Problems in The Design of Materials and Devices for Automated Teaching

ERNST Z. ROTHKOPF

I INTEND to discuss some research questions which arise from attempts to teach by means of automated instructional devices. There will be no pretense either of comprehensiveness or of covering the most crucial . . . issues which are posed by this enthusiasm-arousing technology. Instead, I shall restrict my discussion to problems which arise in connection with two varieties of subject matter—sensory-motor skills and identification learning. The selection of particular problems has been biased to no small degree by what I, from time to time and for a variety of reasons, have thought interesting and seductive. This is particularly true for the section on identification learning.

SENSORY-MOTOR SKILLS

One of the areas in which automated teaching devices offer the most obvious promise is the training of certain sensory-motor skills. This includes such skills as typewriting, the operation of calculators and business machines, stenographic devices, and the like. Training in all of these skills is characterized by prolonged periods of drill which are frequently distasteful for both students and instructors. Progress in acquiring these skills is nearly always measurable in terms of two dimensions, reduction of errors and increases in speed. All of the skills described above involve the operation of apparatus, a typewriter, or a calculator, or some other similar device. The characteristic is unusually favorable for automation of teaching, since the operation of such apparatus can readily be sensed, i.e., be converted into some electrical signal which can be evaluated by the teaching device. In other words, these skills provide a ready-made response language which the teaching machine can understand. Furthermore, the stimuli to which the student must learn to respond are generally well defined and can be readily manipulated by the experimenter in the practice situation. In view of so many favorable conditions, it is surprising to find so little evidence of work in the development of automated teaching devices for instruction in sensory-motor skills.

To the best of my knowledge, the only serious effort in training sensory-motor skills has been that of Gordon Pask and the Solartron Electronics Group in England (Pask, 1958a). One of the problems which they have tackled is the training of operators for the 12-key Hollerith card punch. They have developed a machine, called SAKI, which consists of four components. These include a "conditioned stimulus" presentation gate [1] which presents the numeric stimuli to the student, an "unconditioned stimulus" gate consisting of appropriately coded lights which are arranged in the same spatial pattern as the keyboard and which prompt the student about the location of the button he is to press. The remaining two components are a response gate—the standard Hollerith push-button console—and a computer which includes a performance memory. The events at the stimulus gates are controlled by the computer according to two input (performance) variables from the response gate. These input variables are (a) punching errors and (b) response latency. Pask's training methods are akin to quality control procedures; i.e., rate of presentation of the stimulus numerals is slowed down when error rates are too high, and stimulus presentation is speeded up when error rates are getting too low. In a somewhat analogous manner the intensity of the prompting lamps in the "unconditioned stimulus" gate diminishes as performance improves.

A most interesting research issue here is how far the carrot should be kept from the donkey's nose in order to make him run

---

This paper was prepared while the author was on the faculty of Rensselaer Polytechnic Institute. It is substantially identical to a paper presented at a symposium on "The Automation of Teaching," held at the 1958 meeting of the American Psychological Association, and is based in part on the author's work in 1956-57 on the self-instructional training project initiated at the Maintenance Laboratory of the former United States Air Force Personnel and Training Research Center. Dr. Rothkopf is now at the Bell Telephone Laboratories, Murray Hill, New Jersey.

[1] This term, as well as others which are used to describe machine functions throughout this article, is derived from a general schema for automated teaching situations which was described in an earlier paper (Rothkopf, 1957b). These terms do not imply physical components but rather delineate functions which one or more components of the device or the program may perform.

the fastest for the longest time. How good should performance be before the stimulus presentation rate is speeded up, how far can performance be permitted to deteriorate before the presentation rate should be slowed down? Do the principles which apply in this matter hold for all levels of performance?

The Solartron workers appear to have avoided the general issue but have reached a successful empiric approximation. The simple SAKI computer plays a game with his student which few seem to be able to resist. SAKI races his pupil and the pupil's attempts to catch up make SAKI run even faster. Mr. Pask reports interesting elation effects in connection with these races. Students are said to underestimate the amount of drill time which has passed and, further, some students are reported to practice until they are exhausted. It behooves us to attempt to replicate such findings under carefully controlled laboratory conditions. However, even without the promise of such exotic effects, the relationship between performance, pacing, and learning is at this time one of the most challenging empirical questions for the automation of sensory-motor skills training.

Rate of stimulus presentation in SAKI is controlled by varying the exposure time of the individual stimuli. If performance with a particular symbol, say the numeral "8," has been good, SAKI gives the student less time to respond to 8 the next time that stimulus symbol is presented. Varying the *frequency* with which each symbol is presented in drill, according to performance with each of the various symbols, may seem to be the next logical step in the automation of sensory-motor skill training. This step requires careful empirical appraisal before we burden the performance memory with another expensive feature. Several studies have been carried out in the past in which the frequency of signal presentation in Morse code reception drill was made proportional to signal difficulty (Morsh, Stannard, and Graham, 1948; Sidman, Keller, Kennedy, and Wilson, 1955). These studies indicate that either small or no advantage resulted from this training strategy. It may be that in learning tasks in which the main difficulties arise from stimulus discrimination problems, *the kinds of errors* which are made should be considered in determining stimulus frequency rather than the number of errors per symbol. For example, in Morse code training many errors are made to the letter H but relatively few to the letter S. However, if most H errors involve making the response S to H, both S and H drill frequencies should probably be increased rather than the frequency of H alone.

Another problem which may result from weighting the frequency of certain drill events by their difficulty is the formation of response sets. The work of Estes and his students indicates that subjects tend to match stimulus probabilities with appropriate response probabilities. This tendency may interfere with acquisition of the criterion skill or may give a misleading impression of the quality of performance. A better understanding of stimulus probability matching behavior would be helpful before we recommend how automated training devices should weight stimulus presentation frequencies by stimulus difficulty in repetitive drill.

IDENTIFICATION SKILLS

Another area in which the automation of instruction is likely to be of utility is the training of identification (paired-associate) skills. The identification skills which appear most suitable for automated training are those which involve verbal response terms. Examples of such skills, which pose practical training problems, are: nomenclature training for maintenance personnel; Morse code receiving; lexical language training; synoptic code for meteorologists; other codes such as semaphore or blinker; the identification of aircraft as well as ships, both direct visually or of their representation on radar screens; and such diverse other skills as the identification of birds, fungi, plants, or even speech spectrograms. The general problem of identification skills training appears to me especially important because there are a number of reasons to believe that certain identification skills are propaedeutic to the acquisition of a number of complex conceptual skills. Furthermore, it may not be unreasonable to expect that we may someday learn to analyze some complex conceptual skills into components which can be fitted by the paired-associate paradigm. The behavioral data obtained by use of the memory drum may even survive the not uncommon conceit that "new psychology" will relate it to the dust bin.

Training of practical identification skills with automated teaching devices has been attempted on an experimental basis. The work of Briggs (1958) with the Subject-Matter Trainer is one example. But although identification skills are presumably among the simplest kind of subject matter, their training poses a number of programming and device design problems for which only very inadequate rational solutions exist at present. I would like to discuss some of these problems here.

One of the questions which arises in connection with the training of identification skills is the sequence in which materials are to be programmed. Some of Gagné's work (1950) suggests that for many types of complex stimulus materials, e.g., aircraft identification, nomenclature for mechanical parts and the like, it is good programming strategy to present similar stimuli in immediate sequence. My own work with Morse code reception drill (1958a), on the other hand, indicates that it is more economical to interpose as many dissimilar signals as possible between any pair of highly similar stimuli. There is some suggestion that similar stimuli should be clustered in the drill program sequence when the student has many analytical verbal responses to the associative stimuli in his repertoire prior to the beginning of practice. On the other hand, if stimuli are *not* capable of evoking many analytical verbal responses, then maximal separation of similar stimuli will yield better training results. The notion that a taxonomy of stimuli based on the likelihood of analytical verbal responses has nontrivial training consequences is in sore need of empirical testing. I am at present trying to develop a technique which seems to be needed in such a test; namely, a quantitative, response-inferred index of the likelihood with which stimuli are able to elicit such analytical verbal responses in representative subject populations.

Another source of interesting programming problems in the automation of identification skills training is a number of variants of the classical whole-part question. For example, under what circumstances is it more efficient, from an instructional point of view, to split training of a set of identifications into smaller subsets? The theoretical work of McGeoch (1942) and Orbison (1944) suggests that it should become increasingly more economical to split a set of identifications into smaller training parcels as the number of identifications in the total set becomes larger. This view is supported for two-way splits by an excellent but little-known experiment by Orbison (1944). Empirical verification is required, however, for the extension of Orbison's thesis that the most economical training strategy requires more and more parcels as the total number of identifications in the set increases.

The theoretical approach of Orbison to the whole-part problem was based on the number of discriminations which have to be made in learning to identify a set of stimuli. Furthermore, it was based on the assumption that discrimination difficulty between any two stimuli, in the set of stimuli to be identified, was equal. The latter condition is rarely met in ordinary identification training situations. Similarity among stimuli generally differs widely, and, consequently, difficulty of discrimination among stimulus pairs is quite heterogeneous. It would be most useful to develop and test a rational method for weighting the number of discriminations by a predictor of discrimination difficulty such as, for example, stimulus similarity.

There are other problems related to the whole-part question which are interesting. Once it has been decided to use the part method, to what level should training with any subset be conducted before initiating practice with another subset—or before practicing the entire set as a whole? What variables bear on this decision? Is there any advantage to be derived in carefully controlling the stimulus composition of the various subsets (see Rothkopf, 1958a)? Assuming a given stimulus composition for each of the various subsets, is there any special advantage to be gained from practicing the several subsets in any special sequence (see Rothkopf, 1957a)? It is suggestive for the experimenter that all of the whole-part questions, which are mentioned above, depend for general answer on improved understanding of the relationship between the strength of each of several identification habits, and interference as well as facilitation among these habits. This looks like a condition which is friendly to the growth of a simplifying theoretical structure.

Another, somewhat different, family of unanswered questions stems from the use of the method of adjusted learning in identification training. Concern with this method stems from attempts to incorporate, in automated teaching devices, features which will adjust future practice sequences to the student's past performance. I believe that such features are the reason why we can call the teaching situations which we are discussing *automated* teaching situations. The control of the practice program in terms of the student's past performance is analogous to the operating mode of an automated milling machine or lathe. Industrial machines such as these can adjust the actions of the cutting tools in terms of a preset program *or* in terms of the discrepancy between the surface which is being machined and some preset standard. Similarly, in the automated teaching situation the practice can often be continually controlled in terms of the student's success or in terms of the kind of error he is making. One of the simplest ways of doing the first is through the method of adjusted learning.

The method of adjusted learning, described by Woodworth (1914), consists simply of a practice scheme for controlling the number of times the subject has responded correctly to any item in a set of identifications. Some criterion of "mastery," such as, say, two correct identifications or three correct identifications per stimulus is chosen, and each stimulus is then eliminated from the practice order as soon as the criterion is reached for it. This is one way of automating practice since the composition of the list is continually being adjusted in terms of the student's growing mastery of the list. Two problems suggest themselves here: (a) What is the relationship between the criterion which is chosen and characteristics of the practice materials such as similarity among stimuli and (b) which of two alternative techniques for implementing the criterion for "mastery" should be adopted?

I have very little to say about the first problem although it is by far the more fascinating. With respect to the second problem, there are at least two methods of implementing a given criterion for eliminating an item from the practice series. Take the criterion of two—not necessarily consecutive—correct responses, for example. You could eliminate an item from the practice order as soon as it evokes a correct identification once and then continue practice in this manner until there are finally no items left to be practiced. Then you could repeat the entire procedure once more. The alternative approach is to arrange practice conditions in such a way as not to eliminate any item from practice until it has produced a correct identification twice and to continue this until there are no items left to be practiced. In either case, at the end of practice each item will have evoked two, and only two, correct responses. I performed an experiment to see whether these two techniques differ in training effectiveness—and found that, at least as far as nomenclature training involving unfamiliar stimuli and rather familiar responses is concerned, there are hardly any differences between the two methods. If this result is supported by somewhat more general evidence, it has economic implications for the construction of teaching devices. It suggests to me that one can build machines with relatively short performance memories without sacrificing training effectiveness. The machine needs only to remember whether an item has been responded to correctly once—and not "how many times" any item has been responded to correctly in the past. This is the sort of functional specification which the design engineer can utilize—and I believe could have saved some money in the construction of one or two existing devices.

INSTRUMENTATION FOR RESEARCH

I would like to describe in this final section two pieces of laboratory apparatus which were developed in 1955-56 while I was on the staff of the Air Force Personnel and Training Research Center. My objective in describing this apparatus is not only to show off hardware which has grown in my laboratory. I also want to make some points about research instrumentation in automated teaching.

The first of the two devices, an experimental response gate, is shown in Figure I. I have called it in a frivolous moment the response gate of the *Polymath*—from the Greek roots for the words "many" and "learn." I built it because I was trying to find out what sort of response languages would be efficient for various kinds of instructional materials and also in order to do some work on the error problem. The experimental response gate understands three kinds of response languages—multiple choice, constructed answers, and two-dimensional tracings of a stylus. The multiple-choice feature admits 0-24 choices. The constructed-response language can consist of words, phrases, or numerals up to 29 symbols in length, which are fed into the device by 29 suitably engraved sliders. The third response language involves a blunt plastic stylus which is supported by a tracing arm. This tracing arm permits the student to do such things as follow map features or trace signals through an elaborate circuit diagram and the like. By means of a pantograph-like sensing arrange-

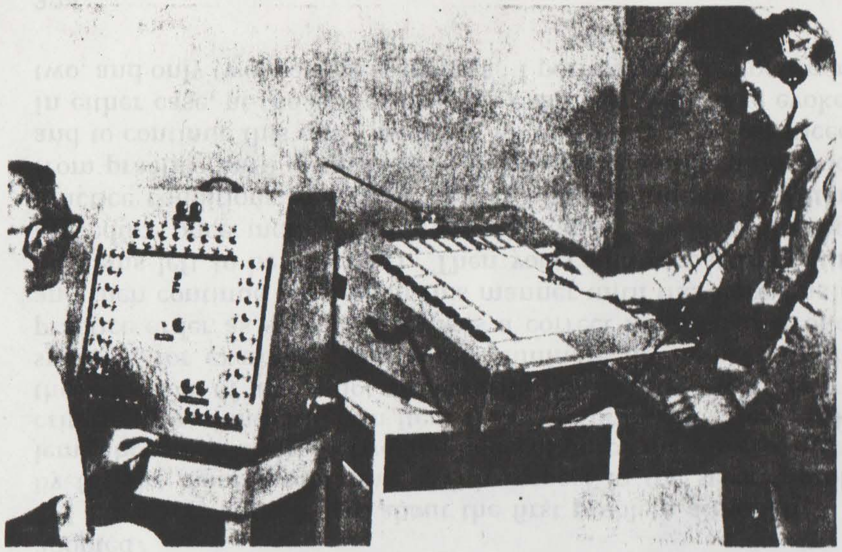


FIG. I. The "Polymath" and associated control equipment

ment, it is possible to supply an auditory error signal whenever the student departs from the correct path. Performance evaluation and feedback for the constructed response and the tracing arm are automatic, in the sense that the stimulus tablet is marked with electrically conductive ink on the side which is hidden from the student. When the student positions the sliders correctly or when the tracing arm is in a tolerated position, an electrical circuit is closed. The device evaluates the student's performance by sensing the state of this circuit and supplies him with an appropriate error signal.

For a number of reasons it was decided not to "automate" the multiple-choice language and a number of associated displays on the student's console. Instead, the student's choice energized a lamp on the experimenter's control console in an adjacent observation room. The experimenter, by means of a rather elaborate array of hand switches, was able to supply the student with several kinds of performance feedback depending on a predetermined set of rules. In a sense, the experimenter acted in this situation as an expensive but readily available computer who performed programming functions. The experimenter's console was designed to help him stay ahead of the student and was provided with a patch board which made it possible to wire correlated events together and so simplify the experimenter's task. I found that this "little man in the box" technique worked well for the problem I tried it on. This problem required the experimenter to vary the number of choices which were available to the student according to performance on the last set of choices, to turn on a "correct response" indicator and also to turn off a lamp under any incorrect choices which the student had made. Simulating programming functions in this manner can save a good deal of time, hardware trouble, and money. After all, the state of the art in electrical and mechanical engineering is considerably ahead of the state of the teaching arts. Our objective should be to supply functional specifications for automated teaching devices and programs and not electromechanical designs. The "little man in the box" technique suits this purpose, particularly as far as identification training is concerned.

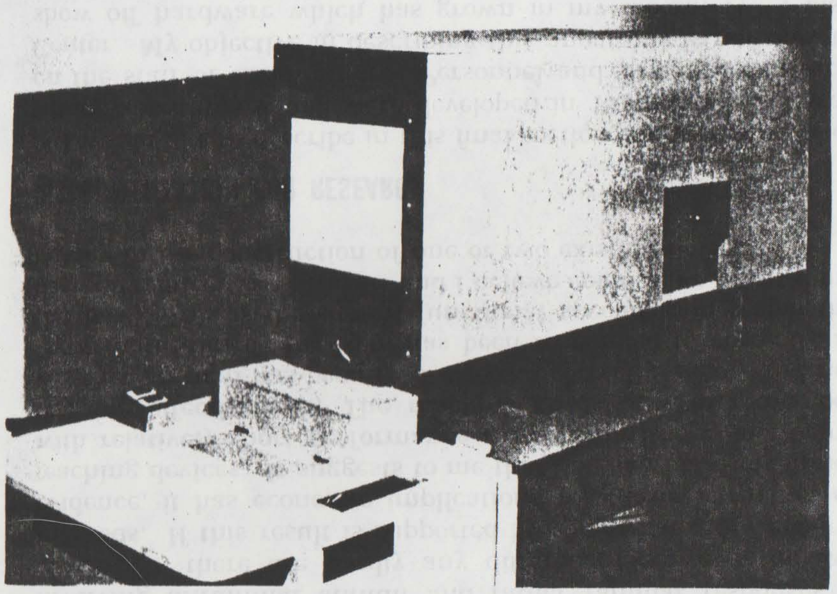


FIG. II. The "Autograph"

"Automated teaching devices" can also often be simulated by common laboratory apparatus. An illustrative example of this kind of configuration is shown in Figure II. It consists of three major components: an automatic slide projector, the paper tape transport of a polygraph, and a Hunter-type decade timer. Problems and answers are projected from the rear on the screen, which is seen in the center of Figure II. The answer is, however, obscured by a shutter located near the plane of the screen. The student writes his answer on the polygraph tape which is exposed to him through a slot in the transparent mask which covers the tape. When the student presses a switch, the timer advances the paper tape a given distance. The student's response is in this way locked under the plastic mask. The operation also shorts the relay which has locked the shutter circuit, and the student can now expose the correct answer. The student scores himself, and this is also recorded on the tape. By depressing another switch, the student can then advance the tape again and also call for another problem slide from the projector.

It is probably clear by now that the philosophy of research instrumentation which I am touting is: Try to avoid constructing sophisticated and expensive teaching machines for research purposes! "Simulating" teaching machines either by use of the "little man in the box" technique or by using common laboratory apparatus will probably do most jobs fairly well.

Finally, I suspect that the information which is most needed in automated teaching technology will come from experiments which do not use any apparatus resembling teaching machines. Those who are interested in promoting the advance of this technology should keep this in mind. We must have a better understanding of the fundamental nature of learning processes if we, as psychologists, are to make useful contributions to the technology of automated instruction.

Problems in Simulation and Programming In the Design of Complex Trainers

LESLIE J. BRIGGS

RESEARCH in automated instruction, up to this time, has concentrated largely upon methods for programming of verbal learning. By "verbal" learning is meant the giving of oral or written responses in learning of paired-associate materials, the memorization of serial materials, the acquisition of facts and concepts, and the application of concepts and rules in continuous discourse material. This emphasis upon programming of verbal learning is of importance for at least three reasons. First, because verbal learning is of great importance to education and to technical training, it is well that we learn how to program such material. Second, since a learning theory eventually must deal not only with acquisition of simple facts but also with the use of concepts in complex situations, programming of conceptual material is an appropriate undertaking. Third, for experimentalists working on very limited budgets, use of the inexpensive devices that are appropriate for programming of verbal learning makes possible meaningful research without large equipment costs.

LEARNING OF JOB ACTIVITIES

With this recognition of the importance of verbal learning per se, I now wish to turn to learning situations in which effective job activities are to be learned. One difference between such job activities and much of the learning material in educational courses is that correct sequences of motor responses represent the criterion performance sought. This is not to imply, however, that the appropriate method of training for all such job tasks would be the direct practice approach used in training for motor skills. Instead, only a few tasks of a job may require direct practice for

---

This article is based on a symposium paper read at the 1959 Annual Meeting of the American Psychological Association. Dr. Briggs is now Director of the Western Office of the American Institute for Research, Santa Barbara, California.

the purpose of acquiring motor skills. Other tasks in which the correct criterion performance also is motor in nature may depend, learning-wise, upon proper application of information and concepts as implicit activity which precedes the motor response. If the reader will permit a very loose term here, I would like to employ the term "mixed" task to refer to job activities in which during successive moments of activity the man may be performing a motor act based on direct practice of a motor skill and then motor acts which are simple in themselves but which are the result of applying a great amount of factual information and concepts to a particular problem encountered in the job environment. I call such a job activity "mixed" because, clearly, appropriate training for the job would not consist entirely of motor skill training, of rote memorization, or of any other "pure" type of learning which we have previously studied in our laboratories.

TYPES OF TASKS IN ELECTRONIC MAINTENANCE

It is my purpose here to suggest that, in training of personnel for many job activities, teaching machines—as we normally think of them—have a place for learning verbal information to support skilled performance. Beyond this, however, I also would like to suggest that special types of teaching machines are needed in which we combine simulation features with self-instructional features to provide both effective learning and learning in a job-like performance situation. To illustrate the questions which the designer of such teaching may encounter, I would like to take the example of design of training equipment for electronic maintenance personnel.

I would like next to outline the nature of some typical tasks required in such jobs and, for each, to raise some questions which a training designer would need to ask himself. Some of these questions involve simulation issues, some involve programming issues, and others involve the question of what the actual learning or training units should be for various tasks and the instructional sequences which might be employed.

Brief Fixed Procedures

Many maintenance activities should be memorized because they are brief, there is a fixed sequence for doing them, and they

are frequently required, thus making it economical to perform them from memory. For such tasks, one might first ask: Would it be more economical to teach the man by direct motor practice of the activity on actual equipment, or would one save time by having the man memorize the task in verbal form and then perform it on the equipment as a sequence of motor responses? Whichever approach is taken, a programming issue arises as to what are the memory-span units or performance segments which should be employed in the learning of such tasks. In any case, there is the need to learn to identify and locate the objects named in the memorized procedure.

Lengthy Fixed Procedures

When a fixed procedure is too long to be memorized economically, job aids in the form of check lists or some other cueing mechanism are provided to enable the man to become familiar with the task during training and to perform it dependably in the field situation. Such job aids for field use, however, cannot normally direct the man in every small step of the procedure, so the training issue is to discover how a man can practice in training with generous cue support but eventually become able to perform the task on the job with the aid of minimum cues. This cue reduction issue may exist whether verbal rehearsal or direct manual practice is the training method. But if a "full cue" job is practical for field use, the man in training often needs only to learn the identification and location of objects referred to in the job aid.

Problem-Solving Tasks

Particularly in the area of troubleshooting electronic equipment, the main training requirement is not to memorize and perform short fixed procedures, nor is it to learn to perform lengthy fixed procedures with the assistance of job aids. Rather, the task is to approach a novel problem situation and to select and employ factual information and concepts efficiently in solving the problem, using variations from fixed procedures as the problem requires. Since there is an almost infinite number of problems—that is—of symptom-malfunction relationships in the equipment, it is not an economical approach in training of troubleshooting to depend upon memory of procedures to isolate all troubles. For this type of task we have not only the question of when to introduce verbal learning and when to introduce direct troubleshooting practice, we also have simulation and programming issues; but, more important, we face questions as to what the unit of training for such tasks should be. Since direct practice in troubleshooting is not the recommended initial training approach, some type of component skill analysis is necessary to identify the major classes of subject matter, concepts, and rules of procedure which the man would need to use to solve such problems.[1]

If the appropriate knowledge and skill components can be identified, verbal training with normal teaching machines might become the training method for a major portion of the subject matter. Also, trainers for developing part-skills are often needed. Even so, eventually during training, practice of troubleshooting in a job-like situation may be required to produce a consolidated, smoothly running total performance and to give the familiarity, or "feel," of the task in a realistic way. The details of what must be simulated and details of programming issues also must be decided.

Thus for three typical kinds of maintenance activities, we raise four kinds of issues; namely, (a) when to employ verbal pretraining preceding simulated direct practice; (b) how to break up tasks into activity and learning units; (c) for complex tasks, how to identify the verbal skill components; and (d) some simulation, programming, and cueing problems.

THE UNITS OF TRAINING

Having described briefly three types of typical maintenance tasks in terms of the criterion performance required and some of the issues they raise, we may now ask more directly: What are the units of training for such tasks? For the brief memorized procedures, if we adopt the verbal pretraining approach first, the units of training appear to be verbal sequences which are effectively memorized when practice is programmed appropriately. This involves memory-span units as well as effectiveness in arranging whole-part practice sequences. Most research in teaching machines has not paid much attention to memorization, and I believe this is a fruitful area for research. If, on the other hand, for brief memorized procedures, we adopt the approach of direct practice upon the equipment, the units of training may pertain not only to the appropriate length of performance segments equivalent to a memory unit, but also to the kinds of stimuli which will elicit the response. If these stimuli do not appear in the field situation, we may have an additional fading or vanishing problem to face. If a fixed procedure is so brief that practice of the total procedure—rather than a part of it—is an effective memory unit, we might say that the task itself becomes a unit of training.

For moderately lengthy procedures, however, we may find that small segments of the total task become the effective training unit. In either case, the training unit is an actual performance segment of the task itself or is a verbal representation of such a performance segment, along with auxiliary cues. One would appear to face more of a transfer problem in the verbal rehearsal approach.

For the very lengthy procedures, which must be prompted by job aids in the criterion situation, we have a similar whole-part unit problem with respect to performance of the unit, but the memory problem is related to the size or coarseness of the cue, which can trigger a short-action sequence required before the next cue is supplied. The identity of the check-list cue and the nature of cues provided by the job environment itself may be important determiners of the appropriate unit of training.

Thus for fixed procedures—brief or lengthy—the task itself, or a portion of it, may become the training unit. For the more lengthy of such tasks, training units may be broken down in a finer manner than the prompting techniques used for the criterion performance; but since direct practice appears to be an effective approach to proceduralized training, dealing with the task itself, or segments of it, appears to be a fairly reasonable basis for designing training.

Turning now to problem-solving tasks, we find that task units themselves are probably not good units of training; that is, the skill and knowledge components required for solving problems appear to be the important things to concentrate upon during training. Thus training for troubleshooting may involve only a little practice in troubleshooting—but it requires much effort in learning needed information, in deciding which information is relevant, and in applying principles and strategies which would be effective in solving the problem at hand. For such a situation, it would appear that learning of factual information, rules, and examples of their application may be the most relevant training units early in the training procedure. Once these are mastered, it may be time to introduce direct practice in troubleshooting, in which each troubleshooting problem is—in a sense—a unit of training. Even so, in prompting the man correctly through such a troubleshooting exercise, no doubt the same facts and rules referred to previously will become the cueing and feedback elements in the training process.

---

[1] It is recognized that efforts are being made to reduce the requirements for this kind of troubleshooting training by development of effective logic diagrams, guides, and automatic fault-isolation devices. The specific learning requirements vary somewhat in these cases.

It may be surprising to some to find how little reference has been made so far to stimuli and responses as the units of training. I am sure that I will be criticized for this, but I also believe that often the applied psychologist experiences some difficulty in analyzing each task to its smallest elements of stimuli and responses while still finding a way to produce, on schedule, the required training and training equipment design. However, there are, no doubt, training tasks in which effective design of training can be practical and produced on schedule and still be as analytical and —if you will—reductionistic as are some of our brief laboratory learning sequences conducted for theoretical purposes. Too often the training equipment designer has designed equipment to provide "realistic simulation" without having conducted sufficient analysis of the tasks to insure that effective cueing and reinforcing processes are employed.

HOW MANY TRAINERS WILL WE HAVE?

It is clear that deciding upon the training equipment requirements for any given maintenance job should involve psychological analysis and also must be practical in application. Clearly, we cannot have a separate teaching machine for each stimulus-response connection to be established, nor can we always take the time to analyze all job segments as minutely as a theorist analyzes a single S-R event. Furthermore, such an analysis, in itself, would not identify the total training procedure insuring adequate performance of the total task. On the other hand, building a simulator which simply looks like the actual equipment and having the trainee practice upon such a device may be wasteful economically and inefficient training-wise. We greatly need to find a method of analysis of real job tasks, which would lead us to discovery of the optimum number and type of trainers from the point of view of economy and training efficiency. Such an ideal form of analysis could lead us into the use of kinds of equipment with which we are not now familiar. But efforts in this direction are not entirely new. In a previous paper there are described three general types of training equipment for maintenance personnel (Briggs, 1959d). This paper describes devices known as Subject-Matter Trainers, which are quite similar to what we are here calling teaching machines. It also describes procedural trainers and troubleshooting trainers of various sorts and raises some specific simulation issues not touched upon here.

I will not speculate further here on the nature of the equipment which might result from the desired psychological analysis. But more effective analysis, no doubt, will involve some of the issues we have raised herein concerning what must be simulated, and when verbal pretraining is an effective approach if followed by certain types of manual practice. Finally, for complex tasks of a problem-solving nature there are difficult problems as to what the component skills are. Firm decisions on the nature of the units of training for complex jobs should contribute to development of special devices which take cognizance of these psychological units of training but which also are sufficiently inexpensive and practical to accomplish the training job required.

Electronic Keyboard Teaching Machines

GORDON PASK

MOST PEOPLE would agree that individual tuition is better than a hard-and-fast training routine. An instructor is flexible in outlook; he changes his method of instruction to suit the individual and leads his trainee along the path which is personally optimum towards proficiency. If you will accept that this private instructor is at an advantage in teaching typewriting, then I need not defend the philosophy behind a teaching machine; for a teaching machine is designed to emulate the ideal instructor. The philosophy behind it is organic control. The controller, in this case, the teaching machine, seeks a compromise with the controlled system, in this case the trainee. The idea is distinct from, even opposed to, the concept of *automatic* control, where the controller is meant to impose a behavior pattern upon the controlled system.

Not long ago the word "machine" implied something authoritarian. Either the machine obeyed a command, like a tabulator or a computer, or it issued a command, like a process controller. Occasionally it was able to do both, according to some built-in, and not always appropriate, rule, like those lifts which visit the other floors first. The science of cybernetics has drawn attention to the fact that many possible machines are not in the authoritarian class, and a teaching machine is essentially a cybernetic arrangement. The teaching machine which we shall discuss has few built-in rules, and very few rules of any kind at the beginning of a training routine. The rules are built up as learning proceeds and are made to reach a compromise with the individual characteristics of the trainee.

As a real-life example, consider the keyboard skill of card punching, where the trainee has to learn a 12-key keyboard, which is operated at the rate of about one depression of a key every third of a second. In the simplest form of training, the

exercises will be different lines of numerals which have to be punched on to the cards. A skill like this is generally taught in class, and an instructor is there to give over-all supervision and to conduct performance tests. However, it is not difficult to imagine what you would do in the position of an instructor if you wanted somebody to learn the skill very quickly indeed. The idea can be extended to more complicated skills, like typewriting, where individual instruction is more common.

An instructor keeps a constant eye upon the trainee's performance. A teaching machine, which simulates an instructor, does exactly the same thing, except that it takes an electrical measure of performance directly from the keyboard, whilst the real-life instructor would look to see what occurred. Like the real-life instructor, it examines how long the trainee takes to respond. It examines how many errors and what sorts of errors are made. It finds out if there are any combinations of numerals or particular jumps on the keyboard which cause more than usual difficulty. As a result of this information, the instructor would vary the sort of exercise presented, so as to suit it to the particular difficulties encountered by the individual. He would encourage the trainee to go faster or slower, according to the trainee's proficiency at any instant, and would point out persistent errors in the performance.

A teaching machine is able to do very much the same as this. It has printed circuits that establish connections which relate the exercise to the keyboard. The teaching machine is informed, via these, whether the trainee is performing the skill correctly. The printed circuit corresponds to the written-out exercise, and both are usually changed about every hour and another set introduced. This takes about two minutes. The differences between its method of conveying information and the instructor's method are quite trivial. An instructor selects an appropriate exercise and hands it to the trainee. Instead of this, the teaching machine has a display which contains several different exercises and is able to switch automatically from line to line, so that the exercise in use at any instant is indicated by a light which comes on. The instructor would probably ask the trainee to hurry up if he wished to accelerate the pace of his performance, or, in the case of typewriting, a gramophone record would be used to ensure rhythmic action. The teaching machine employs a different but more adaptable method. An indicator goes along the line one

---

Reprinted from *Education and Commerce*, July 1958. Mr. Pask is currently at System Research, Ltd., London, England, and has been serving as systems research advisor to the Solartron Group Research Laboratories.

step at a time and tells the trainee which item he should be dealing with.

The pace at which the indicator moves is intimately related to the way in which the trainee has previously dealt with these items. Thus it assumes a rhythm and allows more time to deal with pieces of the exercise which are difficult than it does with parts which are easy. Finally, whilst the instructor would talk about the keys on the keyboard which needed particular attention, the teaching machine brings up a light on the lower illuminated display, which is a replica of the keyboard layout, and points out the key which ought to be pressed for each item of the exercise. Like the instructor's correction, this helpful information can be and is withdrawn as soon as the trainee is becoming proficient. Thus, ultimately, none of these lower display lights appear, because no helpful information is needed when the skill is learned (see Fig. I).

We have reached the stage of advanced logical design of a typewriter teaching machine, and it may be of interest to note the differences between a typewriter teaching machine and a punch-card teaching machine. They are very slight, in principle, though in practice rather different methods are employed. In a typewriter teaching machine the exercises are more varied, since they consist of words and sentences rather than lines of numerals. Because it is easy to memorize the material, about 25 exercise lines must be provided in the teaching machine and the printed circuit is replaced by a punched-tape arrangement run directly from the typewriter mechanism. The keyboard is, of course, a typical typewriter keyboard, and the helpful information display is made as a replica with the rest keys' positions emphasized.

For technical reasons it is more economical to make a machine of this kind if the keyboard is restricted to 32 keys, and we intend, as a compromise, to exclude the numerals in a typewriter teaching machine—thus allowing 26 letters, full stop, comma, semicolon, colon, question mark, and space bar. The shift mechanism is, of course, included, so that the machine must be able to distinguish 64 kinds of response, using the set of 32 different keys.

There are two issues which must be decided by experimental work. The first of these is whether to provide complete helpful information as in a punch-card teaching machine—or whether to provide it in stages. Thus there is a possibility of specifying "left or right hand," "rest key position, one, two, three, or four," and movement from rest key, instead of merely indicating the key which should be pressed. The second issue is the fingering —i.e., which finger is used to press which key. But there is a method which will probably overcome this difficulty.

Of course, the really important issue in either a punch-card machine or a typewriting teaching machine is what goes on between measuring the trainee's performance, and changing the training routine, and its presentation to the individual. It is pretty clear that there is no hard-and-fast rule by means of which we can describe what goes on inside the real-life instructor. An instructor will say that he builds up a mental image of how he thinks that a particular trainee learns the skill, and he tries to modify the training routine into conformity, as far as possible, with this image. But this kind of model, the mental image which he forms, cannot be written down in so many words or expressed in mathematical relationships. If it could be, there would be no training problem in any skill. Nor for that matter would there be any problem of personnel selection, which could all be done by applying a book of rules. But, of course, it cannot, and the best that the instructor can do is to make statements like, "I try to keep up his interest," "I do practice him on the things he finds are hard, but, of course, if this gets tedious, I let him go on a bit . . ." and so on. In other words, the interpretation which we could place on these phrases, like "keeping attention," and "keeping up interest," will depend entirely upon the trainee himself, and the trainee is the unknown in the problem. In fact, we know little of the process which takes place inside either the instructor or the trainee. There is even doubt whether the question "How is it accomplished?" has any meaning, because the wonderful thing about a human brain is that things can be accomplished in many different ways. It would be *surprising* if a brain did *not* use entirely different methods at one instant, and a few moments later, to produce results which are indistinguishable.

But, for the purpose of making a teaching machine, there is no need to know the internal mechanisms, excepting to know that almost any mechanism can be used. I shall try to show the point of this later but will pass over it at the moment in order to examine the outside appearance rather than the inside mechanism of the teaching system.

Teaching is a special sort of conversation directed initially at finding out how the trainee learns and later directed at persuading the trainee to behave in a way that satisfies an employer. In place of words, different kinds of exercise and different kinds of correction are employed. Instead of grammar, there are relationships between the exercise material and the keyboard. Because of this, you may prefer to regard the system as a game. The game analogy is very convenient in any case, because it can be reduced to fairly strict mathematical form and the actual play compared with calculable play. Alternatively, more conveniently perhaps, teaching is like a game which the instructor plays with two objectives in view. But the important point which is brought out by the game analogy is that the instructor is playing with the trainee for several purposes, several objectives. The first of these is to improve the trainee's learning process, to make him learn more quickly, and maybe to make him learn more proficiently. The first objective is associated with competitive play. The instructor is always trying to make life too difficult for the trainee, but the objective can only be secured if he manages to find out something about how the trainee learns, for only in this way can he modify the task which he sets him in an appropriate manner. Thus the second objective of the game is to learn about the trainee; and since there is also a perfectly good way in which the trainee can be said to learn about the instructor, the second objective implies cooperation between the players. It is a process of back-and-forth exchange and interaction between the two. Neither objective has precedence over the other. But if both objectives are satisfied, learning takes place and a skill is acquired.

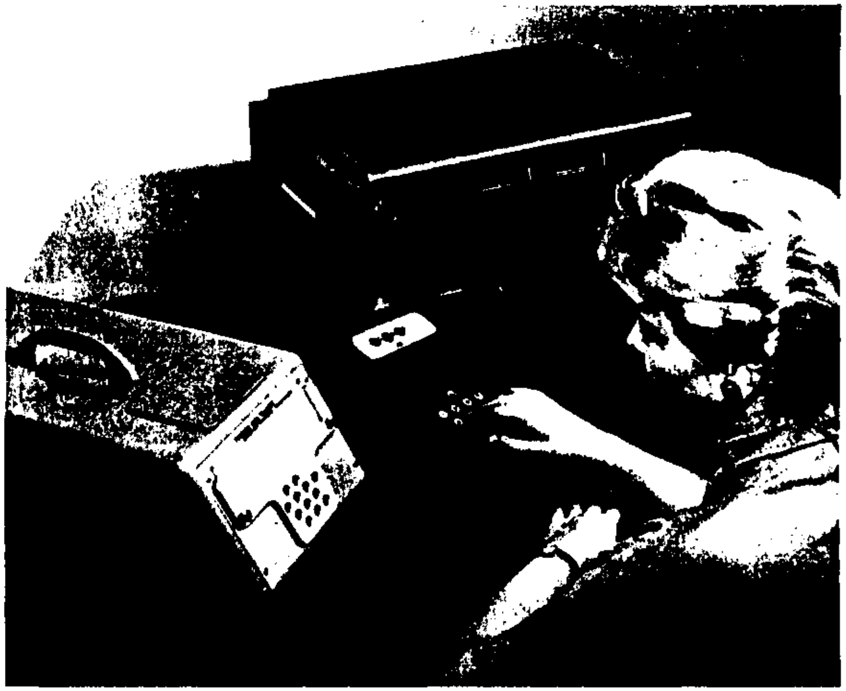


FIG. 1. A keyboard teaching machine (Rheem-Califone). See also Pask 1959b [Appendix I, p. 670].

Simulation of an instructor was a most interesting problem. The first machines were temperamental contrivances, apt to fall apart but able, with luck, to demonstrate the principle. They reflected the flash of insight which tends, for all the theory, to dominate this kind of research. But sometimes a machine has to be manufactured, and at this moment they take away your experimental device and give it to a real engineer.

Fortunately, Mr. Bailey throughout our cooperation has always urged the importance of two things which smooth out the transition from an experimental machine into this sort of arrangement. These two things are a *functional specification*, which asserts what the machine is to achieve, and a *mechanical specification*, which indicates how the machine will achieve it.

The *functional specification* of a teaching machine has been stated already. However, I should like to develop the idea a little. The *functional specification* defines a teaching machine as a device able to play a partly competitive and partly cooperative game with the trainee. The game is to be played in terms of the skill. Thus in punch-card skills the rules of the game, which must be obeyed, are the relationships between items in the exercise and keyboard positions. The different moves which are made by the machine amount to presenting different exercises, and allowing different intervals to deal with the items, and different amounts of helpful information. The moves of the trainee consist of making key-pressing responses on the punch-card keyboard.

To complete the specification, we shall say that the teaching machine at any instant decides about the next move, by analogy with the trainee who decides about his move, namely, about which key to press. Of course, in the case of a teaching machine there is no suggestion of a conscious process. Competitive play means that the moves of the teaching machine render the trainee's life more difficult. He must keep up with an accelerated pace, less helpful data, and more difficult kinds of exercise. Cooperative play reverses the process, and the conditions become less difficult.

Now we have to face the more involved problem of the *mechanical specification*—of saying what kind of mechanism will play the game—in practice. I shall not attempt a technical description. There is no real value in such recipes as "Take 96 condensers, eight relays, and a dozen valves." These components can be broken into a dish and made into a teaching machine in many different ways, and it is hardly appropriate to consider the advantages of each method, because these do not affect the general principle.

Instead, I should like to consider how the game is played, without particular reference to the elements in the machine which play it. To start with, recall two points which were made. The first of these is that the analogy of the game does nothing to describe the process either in a real-life instructor or a teaching machine, which takes place between measuring the trainee's performance and changing the training routine. Further, it has been suggested that we do not need to know about the process in detail, at any rate to make a teaching machine.

The second point to recall is that any arrangement, either an instructor or a teaching machine, must have learned about the trainee in order to know how to play the game. Stated in electrical terms, this means that a model must exist, either inside the instructor or inside the teaching machine, which represents the trainee. Its existence seems to contradict my previous assertion that we do not need to know about the mechanism inside, and indeed there would be a contradiction, except for the fact that the model is of a very special sort. Its characteristics can be asserted concisely:

1. It represents the preferences and oddities of an individual, in particular his idea of what items are difficult.
2. It is, of course, a simplified model.
3. It is always changing.

4. It is built up as learning proceeds rather than built into the teaching machine by its designer.

These properties allow the teaching machine to obey its *functional specification;* further, it can be shown that any teaching machine must contain a model of this kind.

The weakness of the game analogy lies amongst the undefined words such as "difficulty," which are used quite glibly in the description. After all, the only meaning which can be given to "difficulty" is something which this particular trainee finds difficult. There would be little point in building our own idea of difficulty into a teaching machine, and still less an average difficulty scale, for example, a scale obtained by averaging the results of a number of tests using this exercise material presented to different subjects. This average measure of difficulty might be perfectly valid on the average, but it would almost certainly never apply to a specified individual. In fact, even for the same individual, something deemed difficult at one moment will be rated easy at the next.

But the teaching machine must be aware of which items are difficult at any instant and for the particular trainee involved. Otherwise, it is pointless to say the teaching machine plays competitively and makes life *more difficult* or plays cooperatively and makes life *less difficult.* Thus the model in the teaching machine is really a representation of the trainee's ideas of difficulty.

The model is changeful because of the changes in these ideas, which occur when he learns a skill. Since any one item may be more difficult or less difficult than another, it is necessary to measure this value for each item separately. Moreover, since aggregates of items in an exercise may be more or less difficult as a whole, quite regardless of their individual difficulties, it is necessary to measure for each exercise as a complete entity.

For each item in an exercise there is a storage device which registers an electrical potential. Each potential is a measure of individual difficulty accumulated by the trainee. It would, for example, be possible to read off all of these measures at any instant. This would tell us the state of the machine, but, of course, at the next instant its state would almost certainly have changed, because at least some of the potentials would have changed.

However, the fact that at some instant the teaching machine is presenting the trainee a particular item in a particular exercise (and that this item is associated with a certain interval in which he is allowed to deal with it and a certain amount of helpful information delivered by the lower display) is determined by the potentials which exist at the instant concerned, and those which have existed, probably different ones, in the past. The process by which the teaching machine arrives at this state is a decision process. Thus, meaning exactly this, one says that the teaching machine decides upon a move on the basis of a model.

The model is built up as learning proceeds rather than being built in at the start. It is built up by measuring the trainee's idea of difficulty for each item. Since nobody asks the trainee whether he considers an item difficult to deal with or not, how is the measure actually made? The answer is, by balancing his response with an opposite change. At first sight the balancing process looks rather as though it would make the trainee chase his own tail. The reason that this doesn't occur is provided by two facts of psychology, which are vouched for by both experience and a great deal of careful experiment. The first is that there is a maximum rate at which decisions can be made. Beyond this limit, overloading occurs and probably confusion. Thus the teaching machine cannot increase the pace indefinitely, and the limit is imposed initially at any rate, by an ability to make decisions rather than by muscular fatigue.

The second fact is that to teach somebody a skill, you must keep his interest concentrated on the job. This, in itself, recommends the idea of increasing the required pace of decision making. If there are not enough decisions to be made about the job, it is well known that the trainee will be bored. He will think of other things, possibly delightful things, like the summer holiday or the film he saw last night, but these are not strictly relevant to punch-card operating, or typewriting, or whatever is being taught. But the teaching machine cannot interest or fill the attention of the trainee simply by increasing the pace at which decisions have to be made. As we have noticed above, this would result in confusion. The way out of this dilemma appears on examination of the first fact in detail. Thus, it is quite true that a trainee is unable to make more than a certain number of decisions per second. But these decisions may be about almost any entities. In fact, we construct rather complicated entities, and we have to all the time. Otherwise, the detail of the world around us would be utterly confusing.

Suppose, for example, that you are walking across a room with a tea tray and that you wish to put it down somewhere. The ability to do this is, after all, a rather primitive kind of skill. It involves making decisions about alternatives in the world around you. The world as you see it consists of innumerable objects. It consists of Queen Anne chairs, armchairs, upright chairs, chairs with twiddly tops; it consists of occasional tables, dining tables, solid oak tables, mahogany tables, or whatever kind of table you have in the dining room. But when it comes to navigating between the door and the window, which is roughly where you want to put the tea tray, then you don't really think of the room as containing all these entities. Instead you adopt a complicated and very comprehensive entity, "things on the floor"; and when you come to decide, "Do I turn right, left, or go straight ahead?" then you simply decide on the basis of whether or not there is an obstacle, a "thing on the floor," in the way.

It is interesting also that when you come to put the tray down, you have to make another sort of picture of the world with different entities. In this case, probably the two big categories "table" and "chair," because when you are near the window, any object in the category "table" is a place to put the tea tray, whilst any object in the category "chair" is excluded. Thus you get rid of the tray by deciding between them.

Exactly the same comment can be made about a skill, typewriting, punch-card operating, or whatever you wish, except that when you are learning, the categories of objects have to be built up rather than being ready-made. In the case of punch-card operating, the entities which correspond to the detailed specification of the tables and chairs are the separate items. When you are beginning to learn the skill, you conceive of a separate item in the exercise and the particular key which corresponds to it on the keyboard. But, very soon, as you are forced to make decisions at a greater rate, you group these into complex entities, like the categories "table" and "chair," in the illustration; and you must do this to conserve the limited number of decisions per second which are available. The complex entities appear as sequences of items in the exercise and movements on the keyboard which amount to a whole sequence of actions. Instead of being faced with a decision about whether the Item X in the exercise corresponds to the middle or upper left-hand key or to any of the other keys on the keyboard, as you are at the beginning of your training, you are faced with a decision as to whether a sequence of items in the exercise, perhaps XY102, corresponds to one or another of a sequence of actions which you know you can perform on the keyboard.

In brief, you are forced to pattern your behavior. The teaching machine is responsive to this pattern of behavior. Further, it seems reasonable to call this pattern of behavior which you adopt a model, a simplified, changing model of the skill which you build up.

When the teaching machine builds up its model of your behavior, it keeps pace with you and makes you keep pace with it, and in the process it mirrors the model which you are building. Thus the model in the teaching machine comes into correspondence with the model you are building. Because of this and only because of this, it is reasonable to say that the teaching machine is able to play the partly competitive, partly cooperative game as required by its functional specification.

The effect of a teaching machine can be examined either numerically, or by introspection on the part of trainees. Numerically, we expect to obtain an acceleration in the rate of learning of between seven and eight times, for the first hour of a training process. This figure refers to the punch-card skill, but the interval for which this order of increase applies will be longer if the skill is more complicated; for example, in typewriting, one would expect a considerably longer specified interval. The greatly increased rate of learning persists only for this interval. After that —in fact, as soon as the skill begins to be mastered—the increase is less pronounced and drops to a figure of three or four times the usual. This level of learning efficiency persists with the punch-card skill for an interval of about eight hours. Again in the case of typewriting, the specified interval would be a great deal longer. After that a further relative decrease occurs, which is not surprising, since the skill is now largely learned and the trainee is proficient. The term "usual rate" refers to the self-paced learning rate using the same kind of display but without a teaching machine attached.

Some of the more modern learning theories suggest that learning is equivalent to what we have called model-building. It is also the case that, when a trainee is interested, his efficiency in

model-building increases. Thus we might predict that a great deal of the increased rate of learning was due to more efficient model-building activity on the part of the trainee. The prediction is confirmed by introspective reports from various trainees. They report that their interest is kept up and that the teaching machine leads them into the skill along an appropriate path. They express this, of course, in different terms. Some people say that the teaching machine is like a live player in the game. Other people say that it is like a mirror of themselves. At any rate, it is certainly the case that people get annoyed with it and infuriated when they are at a particularly troublesome part of the learning process. But this, surely, argues in favor of the mechanism, if it is a teaching machine rather than a gimmick.

There is another application of teaching machines which deserves mention. This is a personnel selection and aptitude testing. Any teaching machine can be used for this purpose if it is fitted with an arrangement for recording the potentials on the storage devices in the teaching machine as learning proceeds. The recording may be done, with more or less simplification, to suit the kind of prediction needed. Thus in the simplest case a continuous pen recorder and a couple of counters will provide all the necessary data. On the other hand, if you wish to indulge in personnel selection in a serious way, it is advisable to record all of the potentials and thus as accurate as possible an image of the model.

The practical result of using these methods is that a prediction can be made about whether or not a trainee is suitable in between half an hour and two hours, depending upon the keyboard skill concerned. The procedure is simply to use a teaching machine for the first part of the training routine, in which the exercise material is slightly modified to include many different items. The model which the teaching machine builds up is recorded, and either by a mathematical analysis or the application of rules which can be defined after a sample of trainees have been tested in this way, certain features of this model can be said to exist or not to exist. According to the existence or nonexistence of these behavioral features, the trainee is either accepted or termed unsuitable.

There is a basic difference between this and any of the usual methods of aptitude testing. In any other aptitude test a fixed set of exercises is presented, and presented in a fixed manner.

In this case the manner of presentation is determined by the trainee. In other words, an attempt is made to stabilize the relationship between the trainee and the exercise material. The method measures not an arbitrary performance, which is perturbed by all manner of extraneous issues like nervousness and so on, but a way of dealing with a problem. It measures, to be precise, the model which characterizes a player in the game who is just able to beat the trainee. The method is, as yet, in its very early stages but promises to be a new and rather humane method of personnel selection.

I have stressed the close relationship between this teaching machine and a typewriting teaching machine, for it is natural to look beyond the immediate production line of teaching machines for punched-card and keyboard skills into immediate possibilities. But one can look even further and examine the potential applications which are as yet in their laboratory stage. Some years ago McKinnon Wood and I did some experiments which suggested that if this principle worked at all, it would apply to any skill. In practice, the prediction appears to be confirmed. The limits to applying the principles of a teaching system are set by the purely practical issue of obtaining an electrical signal to indicate the trainee's performance.

Some of the applications which work in the laboratory are: (a) the skill of working an adding machine, where the system is modified to deliver groups of items—namely three-, four-, five-, or six-figure numerals—in place of a single-figure item in an exercise line; (b) pursuit skills which introduce a continuous search for a moving object; and (c) the highly abstract decision making encountered by managers and administrators, in which although there is often no correct rule to determine a particular action, there is usually a correct pattern of behavior as a whole.

In conclusion, I suggest that there are very few limits. The teaching machine which we have brought is a representative of a very large class of systems. The physical mechanisms are different, but the over-all principle of a system which is able to learn, to be modified by its environment, and to develop is common to all of them.

The fact and the philosophy of this border country is a new, and to my mind, an exciting field.

Adaptive Teaching With Adaptive Machines

GORDON PASK

INTRODUCTION

IF A MAN is speaking without restriction, I cannot usually predict his next utterance. If two men are speaking together in conversation, I can—for there is a universe of discourse which includes only a finite number of possible statements, related, in some way, to a common topic. In physical terms the conversation is a "closed system"—a system in which all possible changes and events can be enumerated. One of the two men may dominate the conversation. If the path by which he guides his conversation partner from one topic to another conforms to a known curriculum, I should say he was acting as an instructor and teaching a skill. (The word "skill" is used in the wide sense, introduced by Bartlett, to include abstract and conceptual as well as merely motor skills.)

The point of view I shall adopt is that teaching necessarily involves this dominant "conversation partner" and that mental measurement, if it is to be objective, must be made upon the closed system of a "conversation" rather than the open and unpredictable system of man alone. The issue has been elaborated elsewhere by the author (Pask, 1958b). The viewpoint is relevant because the dominant partner need not be a human being. Instead it may be an artifact, usually an electronic device, called an adaptive teaching machine. There is ample evidence to show that such adaptive teaching machines can establish conversational rapport with a human subject. So far, there are only scanty data regarding their efficiency as teachers.

This hitherto unpublished paper has been prepared by Mr. Pask as a summary of some previous technical papers on this subject. With the author's permission, the editors have condensed the original manuscript for inclusion in this volume, deleting some of the more technical portions and introducing other minor editorial changes. Mr. Pask is at present at System Research, Ltd., in London, England.

In short, I believe that teaching machines should be artifacts or automata which are similar in kind to the subject who receives their instruction and that adaptive teaching machines can—at least in principle—sufficiently approximate this requirement.

Some teaching machines are tools, logically comparable to a desk calculator or a visual aid. Others, where a feedback is taken from the subject, are tools which are modified by use. An adaptive teaching machine aims to be more than this—it aims to be an extension of the subject's brain—an adaptable piece of the world in which the subject can exteriorise part of his learning and thinking. This design objective makes them expensive and often unwieldy. Their effectiveness, from a practical point of view, depends upon economic issues. Highly restricted adaptive machines such as the system described in the article "Electronic Keyboard Teaching Machines" (Pask, 1958a) are possibly exceptions to this rule. For the moment, at any rate, the chief virtue of these adaptive machines is that their use offers some escape from utter empiricism, for there is a descriptive framework within which we can discuss the conversational teaching process.

In this paper I shall outline the framework and I hope to convey the flavour of the entire field. First of all, then, we shall examine the way in which interaction between a subject and an adaptive teaching machine (or, in our initial exemplar, between a pair of subjects) is most conveniently described. In doing this, we consider the play of a game. Secondly, we shall relate the game to various real-life teaching systems, in which the type of play characterizes the teaching process. Then we shall outline the logic (though not the mechanical details) of a system which is able to act as a player, i.e., a typical adaptive teaching machine. . . .

REPRESENTING INTERACTION AS A GAME

The situation represented in Figure I has the essential ingredients of a game. Two subjects, A and B, interact with each other through an arbitrary communication channel. To play the game, Subject A selects one of *n* possible response buttons and presses it. As a result, one of the *n* possible events will occur in the vicinity of B. After pondering upon the event, Subject B selects one of his own response buttons, and this action gives rise to an event in the vicinity of A. Each selection by a subject is called a "move," each procedure of the kind I have outlined is a "play" of the game, and each particular pair of moves, one selected by Subject A and the other selected by Subject B, determines an "outcome" of the play.

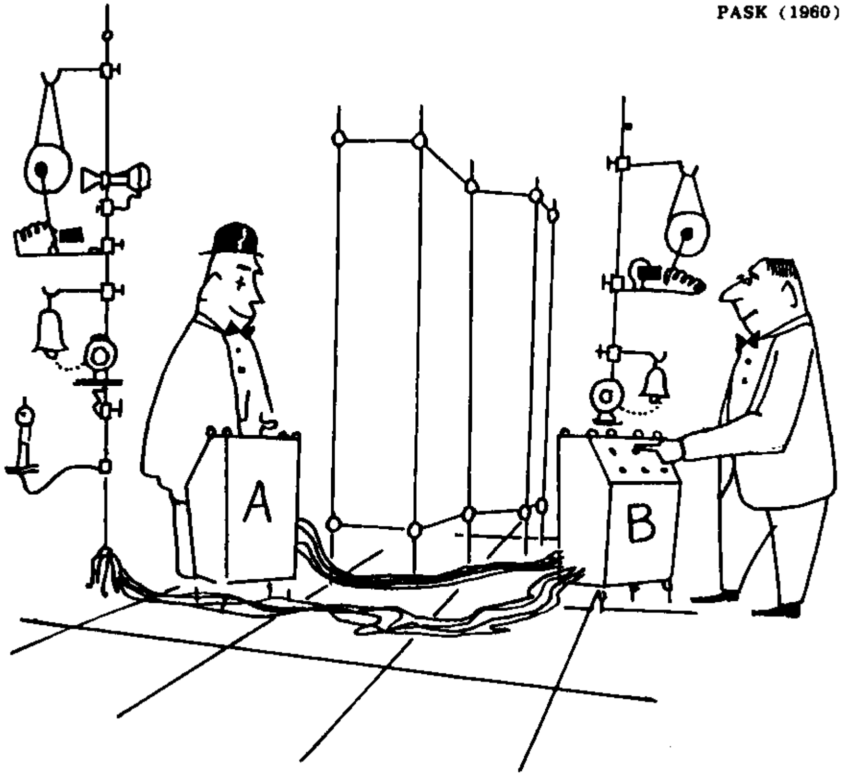


FIG. I. Interaction between Subjects A and B in a two-person game situation.

Play is repeated indefinitely. Because of this the subjects will be apt to think ahead and, instead of deciding about single moves, to decide about and adopt predetermined sequences of moves called "strategies." The word "strategy" will be taken to include single moves as well as these extended sequences, so that we can talk about a more general kind of "outcome" specified by a particular pair of "strategies". . . .

Since the subjects are human beings, it is reasonable to suppose that they will prefer certain outcomes rather than others. Subject A, for example, might prefer the outcome A-Gong, B-

Gong, to A-Bell, B-Bell, and this in turn to either Gong-Bell combination. Subject B, on the other hand, might prefer A-Bell, B-Bell to A-Gong, B-Gong and, in common with A, prefer either of these to the Gong-Bell combinations. It will be convenient, though unnecessary, to suppose that A and B can attach numbers to each outcome as a measure of the relative satisfaction they obtain when it occurs. The assignment of a pair of numbers (namely, an A-Satisfaction and a B-Satisfaction) to each outcome is called a "payoff function." With a preference pattern of this kind, the derived payoff function determines optimum strategies that can be adopted *only* if some measure of cooperation takes place between A and B. Suppose, for example, the payoff function

	A-Gong	A-Bell
B-Gong	(2,1)	(-1,-1)
B-Bell	(-1,-1)	(1,2)

which is about the simplest compatible with the exemplar of preference ordering (the first number in each pair is A-Satisfaction and the second number is B-Satisfaction). The optimum strategy requires collusion, since A must play the same move as B (or vice versa) and "Bell" must be played as often as "Gong." Such a joint strategy secures much better satisfaction than the maximum expected from competitive play when A and B act independently. In general, if some of A's preferences agree with some (but not all) of B's preferences, both players are in a position to gain more satisfaction if there is collusion between them (and this collusion requires interchange of information).

We can assume that the payoff function itself is built up through an initial phase of trial making; that is, collusion between A and B cannot stem from any previous agreement. Rather, as a result of playing the game, they learn about each other and reach an understanding to behave in a certain fashion. The important point, so far as we are concerned, is that *if* the payoff function is of the kind I have indicated above, then the subjects *must* learn about each other in order to maximize their satisfaction.

If Subject A has had previous experience with the system (playing with subjects other than B) and is thus familiar with the outcomes, he is able to dominate the play. In particular, he can restrict B's knowledge of the outcomes—as B tries to investigate them by combining each of his strategies with each of A's—and (without detriment to his own position since he knows the outcomes already) A can play only a subset of his strategies and thus leave many outcomes unobtainable.

This policy need not be so obstructive as it sounds. Indeed, if we make a pair of realistic assumptions, it is possible for A to help B by restricting him. These assumptions are that (a) there are many possible outcomes and (b) a subject has a limited decision-making capacity, because of which, B (faced with the task of investigating so many outcomes) would probably give up. The situation thus is that A would like to play the game, but B does not play because the initial investigation is overwhelming. In order to induce B to play, A restricts himself to a probably unsatisfying subset of strategies (so that now B sees a smaller set of outcomes which he can investigate). After a while, when B has built up a preference ordering over these outcomes, the game can be played. Of course, from A's point of view it is a small and unsatisfactory game; but since at least some game exists, A can introduce fresh strategies and gradually extend the play to include all possible outcomes.

It seems to me that when A adopts this policy (of first insuring that some game exists and then enlarging it to a game he finds profitable himself), he is acting as a teacher. He is a very vague, generalized teacher, of course. He teaches all possible manipulations of the available moves. The interesting point is that the process of teaching is also the only process which, in these conditions, allows him to maximize his "satisfaction."

In the background of any skill there is a collection of events (and of the possible manipulations of these events), for example, the possible ways of hitting the set of typewriter keys. A logical specification of the skill itself is a restriction which allows only a few of these possible manipulations to take place. This definition may seem a trifle strained in certain cases, but it has the virtue of complete generality; for example, it applies equally well to the skill of typewriting and to the skill of learning a foreign language. (However, one should perhaps stress that it defines a skill in the abstract and says nothing about how the skill is performed.)

We can perfectly well define a skill in terms of the odd collection of events depicted in Figure I. The definition corresponds with a restriction saying that only some of the outcomes which *may* occur in the play of the game *will* occur and that some *will not*. Similarly, a teacher (Subject A) who instructs this special skill (rather than the generalised skill which allows any of the outcomes to occur) is one who insists that as a result of the play Subject B builds up a special payoff function. P_A and P_B can represent the payoff functions of Subject A (the teacher) and of Subject B (the pupil).

. . . .

In the framework of this model, teaching implies, in the first place, that the game shall be played—and that some P_B and P_A will develop. Secondly, a restriction is imposed upon the P_B which does develop, and this, in turn, implies a manipulation of P_A. The manipulation can, and usually will, be elaborate. It is not necessarily true, for example, that Subject A plays only the strategies which correspond to real-life contingencies, and it is likely that P_B is critically affected by Subject B's preferences for outcomes which do not appear in the performance of the skill (and which, as onlookers, we should be inclined to deem irrelevant). However, it is true that the development of P_A and P_B must be constantly governed (though not necessarily determined) by the over-all constraints. . . .

Finally, we shall introduce the idea of a score and of a scoring mechanism. The scoring mechanism tallies with a real-life "marking procedure" or "performance test." The score itself, which might be represented on a dial, is a variable which is high in value if, and only if, the logical requirements of the skill are satisfied by "the play of the game." At the outset of a teaching process we should expect the score to be low valued; when learning is complete, high valued. In other words, we can expect few right answers to begin with but more right answers later on.

. . . .

THE RELATION BETWEEN THE GAME AND REAL SYSTEMS

We shall now relate this model to a real-life teaching system. In Figure II assume that both Subject A and Subject B are informed of the score by a scoring mechanism. Subject A, however, has been enclosed in a "black box." So far as Subject B is concerned, this black box looks like a teaching machine which issues forth strategies and receives a measure of what he is doing.

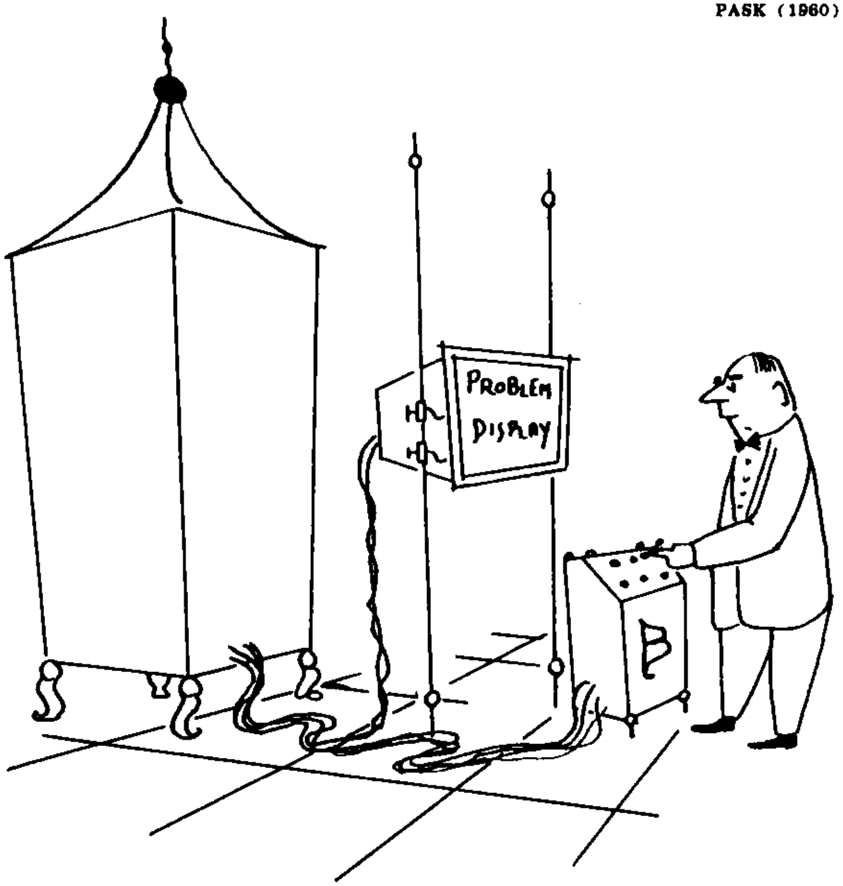


FIG. II. The situation in Figure I with Subject A enclosed in a box to simulate a teaching mechanism.

Later we shall replace Subject A by a real mechanism. For the moment our chief concern is to give a real-life connotation to "moves" and "strategies" and to enlarge a little upon the score. Thus we shirk the task of specifying a machine which plays the game like Subject A and admit that Subject B is facing a teaching machine as phony as Kempelen's Chess Player—just a box with a dwarf inside.

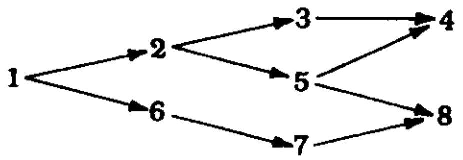
A move on the part of such a black-box teaching machine is any problem which it presents to the trainee, i.e., to Subject B. The calibre of the problem depends upon the skill it is teaching. Thus with conceptual skills in mind, we should encounter problems like "add together 7512 and 8242" presented visually on a card or (as in Crowder's teaching machines) by projection from a filmstrip. The solution to this problem is clearly 15,754, which might be keyed into a response board. Similarly, the presentation, "According to Ohm's Law .. across a resistance of R ohms is equal to R multiplied by the current passing in the circuit" is a problem solved by selecting "voltage" to enter in the blank space from the multiple-choice response alternatives: "power," "amps," "impedance," "reluctance," "voltage," "magnetism," and "myopia."

Another kind of problem is a maze, or its logical equivalent, a puzzle. The subject must find his way out or disentangle the thing, as the case may be. Yet another occurs in connection with fault-detection training where the trainee is confronted with a simulated piece of existing equipment. The teaching machine poses problems by introducing simulated defects which have to be rectified. Finally, it is possible to regard each alphabetic or numerical character presented by a keyboard skill system (such as the SAKI system; see Pask, 1959b) as a problem which can be solved by selecting a particular key.

I would like to stress that a "problem" must be adequately specified before we can speak of a "move." In the case of the fault-detection simulator, for example, we must say: (a) what defect is simulated; (b) whether the trainee is informed of the defect or whether he has to detect it by observing the behavior of the simulator; (c) supposing he is not informed directly, what information does the behavior provide for the subject, on the basis of which he can decide whether the equipment is or is not defective; (d) whether the subject is provided with partial information—for example, whether he is allowed to see inside the simulator and trace back along the causal chain leading up to the defect; (e) how long he is allowed in which to adopt a corrective procedure and thus how long he can take finding out the character of the defect before he tries to put it right.

In the case of the SAKI system, a problem is adequately specified if, and only if, we assert: (a) the character which is presented in the exercise line display, (b) the amount of partial information (in this case, relating to the keyboard locus of the character) which is presented in the lower display, and (c) the interval (which elapses before the next item is presented in the exercise line display) allowed for response.

Thus, in general, the description of a move must include a description of each constraint which the teaching machine imposes upon the subject behavior. A strategy is an unambiguously defined sequence of moves; in other words, it is a sequence obtained by selecting moves from a numbered set of moves according to a rule which states for each move selected one and only one "next selection." This definition of strategy clearly includes the case of a "one-move" strategy, namely, a move. An unbranched teaching machine program is a strategy. Similarly, a branching program can be dissected into unbranched sequences, which are strategies. Thus let certain moves be numbered 1, 2, . . ., 8 in the branched program shown below.



This can be dissected into the strategies: 1,2,3,4; 1,2,5,4; 1,2,5,8; and 1,6,7,8.

If the moves have been adequately specified in the first place, it is always possible to represent strategies in terms of this kind of picture, and we shall assume this possibility in the body of the present discussion. However, it will be worth looking at the pitfalls which beset the identification of real programs with strategies, and many of these pitfalls appear when we try to describe the strategies of a keyboard teaching system such as the SAKI system. The SAKI machine is able to select one of four different exercise lines, each consisting of 24 characters (and it makes a decision to select one or another of these at the end of each line). It is tempting to suppose that the machine has but four strategies to its credit. In fact, there are four *subsets* of strategies specified by the exercise lines. But a *move* is defined by an amount of partial information and an allowed response interval in addition to a character. These parameters of a move depend upon the charges accumulated by $4 \times 24 = 96$ storage devices, one of which is associated with each position on each exercise line. Since strategies are defined in terms of moves and since these charge values fluctuate according to the subjects' performance (and in doing so, modify the moves), there are many more than four different strategies; the exact number depends upon the least difference in partial information and allowed response interval which the subject can discriminate. It is, of course, possible to make a list of each combination of parameter values and thus to define a set of strategies. In the picture we use—which is logically convenient though mechanically misleading—the teaching machine makes an appropriately restricted selection from this set.

By analogy with our discussion of Subject A (in the black-box teaching machine), the strategies of Subject B (the trainee) are response sequences. . . . Responses will count as moves in the game if, and only if, their implication is in some way "appropriate"—and here we cannot assume that preliminary common sense has rendered the whole problem trivial.

I shall exemplify, rather than define, what I mean by appropriate. Suppose that the Subject B is learning typewriting. Subject A in the black-box teaching machine issues typewriting exercises—it will be convenient to suppose he writes them himself. Subject B starts off by making single moves, decisions about single letters in the exercise material. At this stage it would be appropriate if the man in the teaching machine modified the exercises letter-wise, for example, by introducing new digraphs or trigraphs. Later, Subject B will learn to type words, and at this stage it would be "appropriate" if Subject A modified the material word-wise and played sentence-length strategies.

The argument applies with equal cogency to conceptual skills. Let Subject A draw mazes and, after drawing, present them to Subject B. Each maze he draws will embody some logical principle c_1, c_2, . . . and the largest maze he is able to draw, a principle C which logically includes c_1, c_2, ... These mazes, then, are teaching-machine strategies. Subject B attempts to solve the mazes, his response buttons indicating the alternative he selects at each choice point in a maze. Further, assume that Subject B can only solve these mazes efficiently—when presented with many different exemplars embodying the principle C—if he has grasped C. In this case, at a stage when Subject B decides about a sequence of moves sufficient to solve a maze embodying c_1 or c_2, it would be appropriate if Subject A manufactured this length of maze (that is, used this length of strategy), and at a later stage when Subject B can deal with C itself it would be appropriate if Subject A issued the biggest maze he can. As already pointed out, we can perfectly well regard the strategies—exercises, mazes, or whatever they are—as being enumerated and forming a set from which Subject A makes selections, and we are bound to adopt this point of view whenever we discuss the payoff function. The point is, unless Subject A takes the appropriate action . . . Subject B will find this black-box teaching machine either incomprehensible or utterly tedious, and he will simply take no notice of it.

The issue of preserving an appropriate relation between the strategies of the two subjects has already been mentioned in a slightly different guise, for we noted that the subjects must play not only to maximize their payoff, but also to make the game, as such, survive. In a sense, they must speak the same language (a language which changes as they learn), or the game is a fiction. The issue will appear again, a little later (when Subject A is replaced by a machine). There we shall be dealing with the identical problem of how a machine can interest the trainee—and keep his attention.

Finally, we must now examine the scoring mechanism. For some of the skills we have discussed (keyboard skills and skills taught by a didactic procedure) this amounts to a simple comparator. For others, such as the skill of fault detection, the scoring mechanism may be a small computing system. In the simplest case, it is sufficient if, for each move made by the teaching machine, one of n response alternatives is designated a correct response. The scoring device compares the designated correct response (which is selected when the teaching machine makes a move) with the subject's response selection. If the subject chooses the correct response, the score equals 1; if not, the score equals 0. The whole device can be very naive—for example, if questions are presented on cards, each card can bear n separate positions in any one of which a hole can be punched. These positions are sensed by electrical contacts, wired in 1-to-1 correspondence with switches actuated by the response buttons. A hole is punched in the position corresponding to the correct response for the question concerned. The score equals 1 only if the button related to this hole is actually pressed, since only in this case is an electrical circuit completed through the card and the response board. Devices which are similar in principle have been used in connection with keyboard skills.

The scoring procedure can be elaborated in two distinct ways. First of all, it may be impossible to say that a move is or is not correct, it may only be possible to talk about sequences, or strategies as being correct or otherwise. In principle, at least, the teaching machine can inform the scoring mechanism about the strategy which *it* has selected, but the subject's strategy is not revealed until all of the moves have been made. In this case, the scoring device must retain an indication of each move until the sequence is completed in order to compare the two sequences or strategies.

The second elaboration concerns degrees of rightness and wrongness. In the simplest case, we assumed that one answer was right and the other answer was equally wrong. More generally, some answers will be right; some answers, partly right; and some, just stupid. These degrees of rightness and wrongness can be sensed by a mechanism which yields, at its output, a multivalued score variable. Again, there are occasions when we can usefully distinguish categories of mistake; for example, in the hypothetical case on p. 356 the mistaken responses "reluctance" and "impedance" might be placed in the same category, whereas "impedance" and "myopia" are clearly in different categories. In general, there are often logically distinct ways of solving or not solving a problem. If the scoring device is to detect these differences—as it easily can—the score is a set or "vector" of variables, one for each category, rather than the single variable we have assumed so far. Finally, in managerial and navigation training, for example, there is often no unique correct strategy. Since many more or less "correct strategies" are acceptable, the scoring mechanism becomes a device which sets boundary conditions upon the behavior.

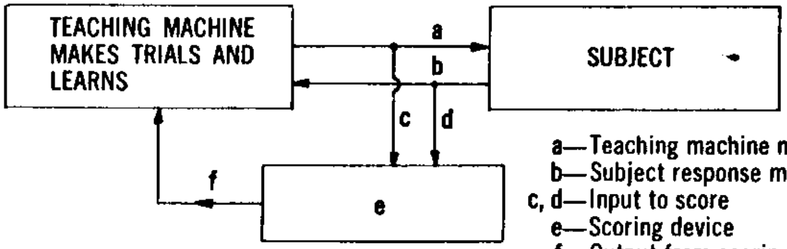
In each case, however, the score is something which refers to the ultimate performance requirements and is in no sense a teacher's evaluation of the performance at each stage in learning. Thus a score of this kind is necessarily computable whenever it is possible to define a skill. The score is fed directly to the black-box teaching machine. It is convenient to think of an electrical connection. Here the score is *turned into* a performance evaluation, either by Subject A or an artifact. Subject B may receive the score either directly or indirectly. In fault detection, managerial or problem-solution training Subject B receives an indication of the average score in the display (commonly on a meter). In less elaborate systems a signal light indicates simple knowledge of immediate results (correct or not-correct response indication). In motor skill training systems, such as the keyboard systems, it is usual to provide some immediate knowledge-of-results data but to omit the score meter since the pace of the presentation, being closely related to the score average, is an indirect indication of Subject B's performance.

SPECIFICATION OF A PLAYER

We shall examine the logical outline of a decision-making device able to play the game and to take the place of Subject A as an instructor. The punch-card teaching machine which has occasionally been used as an exemplar is *not* typical of such a device. True, this machine *does* stem from the same logical stock. But it was bred for a commercial market, and its makeup includes a number of constraints. These are rightly part of a machine which must be reliable and fairly inexpensive. They are introduced because we know, vaguely, how people learn to punch cards. But these constraints cloud the present issue for, in general, we can assume almost nothing about the fitful moods of Subject B. The machine I shall describe is suitable when we, as its designers, are ignorant to this extent.

At the outset we noticed that this game is not always played in an arbitrary interaction between a man and his surroundings. The surroundings must be of a special kind (such as another man) which catch his attention and keep his interest. If they are of this special kind (in particular, if a mechanism can embody this specialty), the interacting members form a closed system in the sense that an outside observer can enumerate its possible states. This notion of closure corresponds to the mechanical notion of stability—if closure is achieved in the system of both members, we can say that one member is stable with respect to the other. Stability does not, in this case, imply a static equilibrium. The outside observer could not say of a stable teaching system that it is always in one state—rather, he might make statistical assertions about it or predict that it would change through a sequence of states. The mechanism which replaces Subject A must first of all create stability of this kind, such that an observer could make consistent, albeit statistical, comments about the system. In order to do this, it must make trials which stimulate Subject B. It must continue to make trials, ever more diverse, until stability is achieved—thus, like Subject B himself, it is instable on its own—stable only in combination with its like. In order to sense stability, it must appreciate the actions of Subject B in response to its trial stimuli. Finally, since we cannot embed within it an image of Subject B, this mechanism must learn by experience how Subject B behaves and fashion a likeness of him in its substance. If this inbuilt model modifies the trials, we can show that the system as a whole will approach a stable condition.

This mechanical picture so resembles the unrestricted play of the game that I have labeled Figure III accordingly—a trial stimulus as a move from Subject A and a response as a move from Subject B. . . . To complete the picture, however, the play must be specific. Thus, in mechanical terms, we must have a system which tends toward a particular stable condition chosen by the outside observer and defined in one of his statements, rather than [just] *any* stable state which occurs. To secure this, notice that the system tends to stability only so far as the mechanism *learns*. Specific stability is achieved by a constraint upon the mechanism, such that it is allowed to learn if, and only if, it is tending to the desired stability. In Figure III we represent this constraint in terms of the game. It is a feedback connection which regulates the permitted learning according to the value of the score.



a—Teaching machine moves
b—Subject response moves
c, d—Input to score
e—Scoring device
f—Output from scoring device gives permission to learn only if score is high-valued

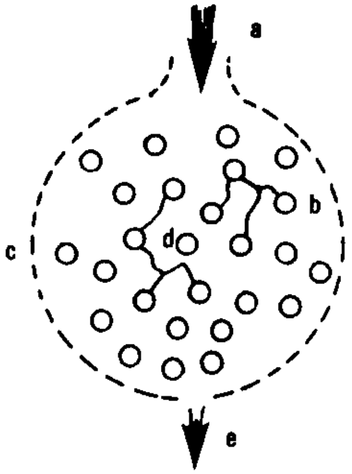
FIG. III. Diagram of interaction between subject and teaching-machine functions

THE LOGIC OF AN ADAPTIVE MACHINE

We are apt to think of a machine as somewhat like a motorcar engine. Different parts have definite functions assigned to them (for example, "the cylinder" where vapor is accumulated or "the spark plug" which ignites the vapor). Thus the names "cylinder" and "spark plug" designate, equivalently, certain pieces of metal and the functions they perform. If the machine is intended to make decisions rather than propel a motor car, we call it a computer; but the same structural comments apply. Inside it, there are valves, condensers, and other components which perform specific tasks which are part of making the decision.

When talking about machines which learn, in the required and nontrivial sense, we must discard this structural image completely. The machines we shall examine must not be thought of as bags full of well-defined components. They are bags full of raw material—if you prefer it, of undifferentiated elements—which may ultimately assume a specific function but which are initially homogeneous.

As indicated in Figure IV, energy flows from an external source into the bag of elements. Energy flows out of the bag into an



a—Energy inflow
b, d—Connections between elements
c—Bag of elements
e—Energy outflow

FIG. IV. Energy-flow diagram

external sink. Each element is credited with one simple property. That is, each element is able to store energy; and if it has accumulated enough, to convert it into a structural connection. Thus, each element is able, within limits, to capture some of the energy in Figure IV and use it to build up a structure. (Further, we can think of these elements as differentiating and acquiring a function insofar as they build up structural relations to one another).

It is convenient—and in one kind of physical artifact to which this model corresponds, also accurate—to say that the act of energy conversion generates a signal. Suppose, then, that a given element generates a signal at some instant; this is conveyed along structural connections to a number of other elements. The paths which it travels depend, in this case, upon any previous activity, for this has built up the structural connections that exist at the moment concerned. The signal affects the connected elements by changing the rule according to which they convert energy—and thus, in turn, it affects the generation of further signals. Four conditions apply to the structures which are built up: (a) A unit of structure has a given energetic cost. (b) Structures do not persist indefinitely, but decay. To prevent their decaying, further energy must be expended; i.e., signals must be sent. Thus there is a "maintenance cost" per unit of structure. But the maintenance cost is less than the construction cost. (c) Thus although structures ultimately decay, they will persist for a short interval. (d) From this brief outline of the activity it is clear that the structures which *will be* built up depend upon the energy available and the structures which *have been* built up previously. Thus if the energy is available, structures tend to perpetuate.

Given some energy inflow, some elements become active, and some structure is produced. We prevent this structure's perpetuating itself indefinitely by restricting the energy inflow. As a result of this restriction, elements in different regions of the bag necessarily compete for the limited energy available. Thus consider the initial conditions, when no structure has been built up; at this stage the elements act independently, and competitively. However, it is possible to so restrict the energy supply that an element can capture more energy if it correlates its activity with other elements than it would if it were an independent entity (correlation, of course, implies that the elements concerned are related in a connective structure). If this is so, the requisite connective structure will appear (and, in general, a sequence of connective structures will evolve) to take advantage of the underlying energetic constraint. An undifferentiated system which evolves in this way is called a "self-organizing" system and, amongst other tricks, is capable of nontrivial "learning."

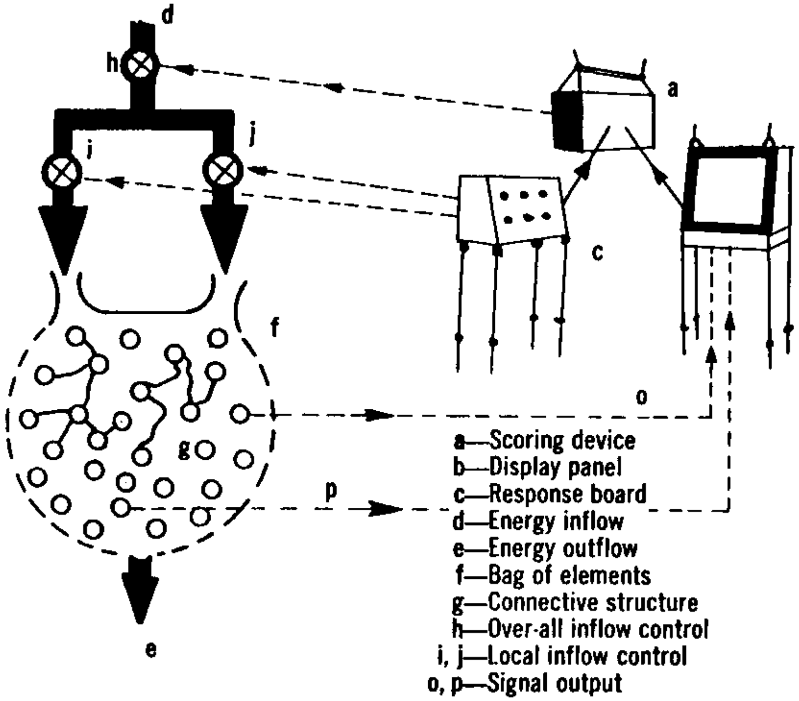


FIG. V. Diagram of an adaptive teaching system

The self-organizing system can be associated with Subject B in a teaching system as shown in Figure V. Subject B's moves act—via transforming devices which I have shown merely as valves in Figure V—upon the energy supply. Signals passing in the system are used—after transformation to select "Subject A," i.e., "teaching machine" moves. The score variable determines the over-all rate at which energy enters the system. The system is able to learn, indeed it cannot avoid learning, due to the various perpetuative characteristics we have discussed, and the Subject-A "payoff function" for the game is directly related to the internal energy distribution in the system.

Artifacts of this kind are not, strictly speaking, finite automata because the system we refer to—when, for example, we say that "the system is learning"—is some active region amongst the bag of elements, not the bag of elements itself. The complete teaching machine is such a bag of elements, together with the appurtenances needed to relate it to the trainee. Such machines have been constructed and used in the laboratory. The more constrained devices have been derived from them.

So far, it has been convenient to construct the "elements" out of electronic parts, and the machine has the appearances of a computer. However, work upon other applications of the self-organizing system led to electrochemical artifacts which may, in a few years, be practical and far less cumbersome teaching machines.

. . . .

[In the concluding portion of his original manuscript, Mr. Pask discussed the relevance of the above-discussed concepts for a possible model of brain function, arguing that "there is some substance to the claim that these adaptive teaching machines are not unlike the people they instruct." This discussion has been deleted by the editors of this volume, since, although interesting, it does not appear directly relevant to the purposes of the present volume.]

A New Technique of Education

SIMON RAMO

WE ARE in rapid transition today to a new world which threatens to be dominated by technological advance. In that new world, (a) man will have learned so much about nature's store of energy and its release that he will have the ability to virtually destroy civilization; (b) production, communications, and transportation will all be "automatic"—these operations of man's material world will have become so vast and complex that they will have to proceed with a minimum of participation by man, his muscles, brains, and senses; and (c) man will conquer space.

There seems little question that these three factors will have dominant effects in the coming decades. The effects are already being felt. (There is even a serious note in the facetious thought that any man who has the courage to stand up and claim that the replacement of man's brains will not have a very important effect on society takes the risk of having his brains among the first to be replaced.)

By hindsight, what is happening would appear to have been entirely predictable centuries ago. We might consider all of man's history until now arbitrarily broken up into two periods, the second of which is now in transition to the new and third one.

In the first period, which might be called the "pre-science" era, man was not consciously employing science to alter his society. Of course, he was aware of the world about him, and he sought to adapt himself to nature's laws. But when he objectively began to develop organized ways of thought and experimentation to further his knowledge of nature, then he entered the era of the "discovery of science and its utilization." One step in this second period is the so-called Industrial Revolution.

In this second era, now coming to an end, man learned to communicate and navigate, to create and harness sources of power well beyond his own muscles, all with the natural result of increased production and fast transportation. In such an era man would be expected to learn gradually more and more about

---

Reprinted from *Engineering and Science Monthly*, October 1957, Vol. 21. Dr. Ramo is currently executive vice-president of Thompson Ramo Wooldridge, Los Angeles, California.

matter and energy, so as to make possible the release of larger and larger amounts, to the eventual point where he could quickly destroy civilization. He would be expected to so increase and speed up production, transportation, and communications as to tie the entire globe together in an enormous network of automatic machinery, cables, and moving vehicles. In time he would have such potentialities in the control and generation of energy and in communications and transportation as to make possible intercontinental ballistic missiles, space satellites, and rockets to the moon. So, finally, he would spread out from the surface of the earth to all the space surrounding it.

And the transition is happening just this way to the new technical age. What makes the present period singular is that now, for the first time, a sharp coincidence between the needs and the state of the art exists. The requirements for the new society, the pressure, and the strength of resources to bring it about now match a sufficiently deep understanding of the laws of nature to make these big steps practical.

CONSIDERING THE NEEDS

Consider first the needs. As one example, note that we are already at the point in air transportation where we badly need a major entry of automatic techniques in the over-all system for airline navigation and traffic control. Today we are alarmed about occasional accidents. But it will hardly take very much increase in the number of cities originating plane flights, or in the amount of air freight to be handled, or in the spectrum of speeds from the very low-speed helicopters to the high-speed jets, or in the demand for reliable operations in all-weather conditions before we approach near-chaos at the airports and intolerable dangers in the air lanes.

Similarly, the military situation demands the extension, and often the replacement, of man's brains and senses by automatic gadgetry. The guided missile is taking over in many areas from manned airplanes. Getting up into the skies swiftly in response to detection of enemy bombardment vehicles, finding the location and the expected future path of such enemy vehicles, making the decisions as to what path should be chosen to effect an interception—these functions, to be accomplished with tremendous rapidity for fast vehicles in all weather, and involving a great deal of mathematical complexity in the decisions, are well beyond the pilot's eyes, ears, and brain.

Business and industry—whether it be department stores, insurance companies, banks, railroads, or ordinary factories—have become so large and complex that the sound containment of their operations from an organization standpoint becomes more and more impossible without the use of automatic techniques.

Now the status of science is such that, without a single new discovery—without even one more new law's being discovered by an Einstein concerning the nature of the universe or the secrets of the nucleus—we can set out to design mechanical and electronic systems that will take over, and do better, many of the things we now do with our brains and senses.

THE COMING CRISIS IN EDUCATION

This rapid and potentially dislocating scientific advance can be expected to heighten . . . the coming crisis in education. Already, the increasingly technical world uses more scientists and engineers, yet the very industrial development that is part of the growing technical society takes the engineers and scientists away from the university and high-school facilities, and the fast world in which we live makes the long study of science seem unattrative to the youngsters. The technical society is complex, rapid, and increasingly dangerous. We can blow up the whole world, yet such a premium is put on the use of our human and physical resources for everything but education that it seems that the new technical society is going to be accompanied by a weakened ability to keep pace education-wise.

Now, if the world were in transition to something different on a very, very slow scale, we could argue that these factors would take care of themselves. Supply and demand would then presumably set to work to make the teaching profession pay off better. Further, the new technical society would be expected to cease to develop rapidly if there were not enough engineers and scientists to make possible that development. So the pace would have to adjust itself to one that would allow all the factors to settle into their respective permanent relative magnitudes.

LOOKING AHEAD

But we are moving much too rapidly for that, and our technical growth is paralleled by social maladjustments still left over from previous eras. The adjustments, instead of being slow and stabilizing, could be chaotic. Obviously, something new is needed. Education should be at the head of the list for priority attention. Our hope for attaining any kind of stability in the highly technical world ahead must rest on the ability to look ahead, to understand the world, and to adjust to it. We must reject such solutions as that we do indeed blow ourselves up; and we cannot accept something approaching a robot-controlled world that consists largely of ignorant and uneducated masses who are slaves to a few individuals who push all the buttons on the machines.

A NEW TECHNIQUE OF EDUCATION

I should like to propose that these very technological advances, about which we normally speak when we talk about the new technical society, must include advances in the field of education, and it is part of the obligation of those of us who are engaged in the engineering side of modern science somehow to apply ourselves to help the process of education. What is needed is a technique of education which is in keeping with the world ahead.

Picture this new technical society—in which the entire airline system, from reservations to blind landing and take-off, is done almost automatically, with the pilot going along only for the ride; in which money is used only in the country communities, and when we buy something in a store we simply put our thumbs up against a little window, our fingerprints are automatically scanned against our balance, and the proper change is made in the respective accounts of the customer and the store—and ask: what will a high school look like at that time?

We have a choice here of two ways to discuss this. We could take it in very gradual steps, starting with the popular suggestion of greater use of television as a teaching aid, or we can allow our imaginations to open up. Let us accept the risk of poor accuracy in prediction and even the risk of exaggeration in order to make a point. We shall describe a technically feasible, even though in some other ways perhaps unacceptable, "modern" high school of the future. But in doing this, please remember that I am neither predicting nor recommending the school I am about to describe, but only using it as a vehicle for making some points later.

SCHOOL DAYS

First of all, we will get the student registered. I won't burden you with the details here; when the registration is complete and the course of study suitable for that individual has been determined, the student receives a specially stamped small plate about the size of a "charga-plate," which identifies both him and his program. (If this proves too burdensome for the student, who will be required to have the plate with him most of the time, then we may spend a little more money on the installation and go directly to the fingerprint system.)

When this plate is introduced at any time into an appropriate large data and analysis machine near the principal's office, and if the right levers are pulled by its operator, the entire record and progress of this student will immediately be made available. As a matter of fact, after completing his registration, the student introduces his plate into one machine on the way out, which quickly prints out some tailored information so that he knows where he should go at various times of the day and anything else that is expected of him.

A typical school day will consist of a number of sessions, some of which are spent, as now, in rooms with other students and a teacher and some of which are spent with a machine. Sometimes a human operator is present with the machine, and sometimes not.

A FUNDAMENTAL LIMITATION

One thing needs to be said at the outset. Any attempt to extend the teaching staff with any kind of mechanical aids would appear to have at least one very fundamental limitation. It would seem that, unless a highly intelligent, trained, and authoritative teacher is available, there is no equivalent way of adapting the material to be presented to the individual student's need, or to judge the understanding and reception of the material and adjust it to the student during the presentation, to discover his questions, weaknesses and misunderstandings, nip them in the bud, and otherwise provide the feedback and interaction between teacher and student that are so essential in transferring knowledge from one person to another.

It is for this apparent reason that, although we can use motion pictures and television to replace a lecturer and can, in theory at least, be more efficient in the use of one skilled teacher's time, enabling him to reach a larger audience, we can only use such techniques for a limited fraction of the total school day. However, you will see in the systems that I propose that, in principle at least, modern technology can go a long way toward removing this apparently fundamental limitation. The whole objective of everything that I will describe is to raise the teacher to a higher level in his contribution to the teaching process and to remove from his duties that kind of effort which does not use the teacher's skill to the fullest.

Let us follow a student who is including in his schedule a course in trigonometry. He will spend a few hours a week on this study in automated classrooms. In the case of trigonometry, only a small part of his time need be spent with a human teacher. Some of his classroom exercises will involve presentation of basic concepts in trigonometry in the company of other students in short lectures, delivered by a special sound motion picture, which uses some human actors who enunciate or narrate the principles to the accompaniment of various and sundry fixed and animated geometrical diagrams.

PUSH-BUTTON CLASSES

However, this classroom has some special equipment. Each chair includes a special set of push buttons and, of course, that constant slot into which the student places his identification plate. The plate automatically records his presence at that class, and it connects his push buttons with the master records machine.

If the class is large, our student is much less likely to sleep or look out of the window than in a normal lecture by a human teacher, because, throughout the motion picture that presents some phase of the fundamentals of trigonometry, he is called upon to respond by pushing various keys. He is asked questions about the material just presented, usually in the form of alternatives. Sometimes he is told that the concept will be repeated and the questions re-asked, this time for the record. He may even be asked whether, in his opinion, he understood what was being presented.

SPECIAL HANDLING

In other words, he is in constant touch with the "teacher"; but something else equally important needs now to be added. His progress and score are used by the electronic master scheduling device to prepare for the special handling of that student in the other portions of the trigonometry course.

At certain other periods during the week, this student continues his trigonometry instruction in a different kind of environment. This time he is seated in front of a special machine, again with a special animated film and a keyboard, but he is now alone and he knows that *this* machine is much more interested in his individual requirements. It is already set up in consideration of his special needs. It is ready to go fast if he is fast, slow if he is slow. It will considerably repeat what he has missed before and will gloss over what he has proven he knows well. This machine continues the presentation of some principles and asks for answers to determine understandings. Based upon the student's immediate answer, it may repeat or go on to the next principle.

With some hints and assistance by the lecturer in the movie, and with appropriate pauses (not accompanied by a commercial), the student is allowed a period for undisturbed contemplative thought before registering his answer.

THE MACHINE'S JOB

This machine is prepared to take a single principle and go over it time after time if necessary, altering the presentation perhaps with additional detail, perhaps trying another and still another way of looking at it, hoping to succeed in obtaining from the student answers that will indicate that the principle is reasonably well understood before it goes on to the next one.

Before the student receives the material from this machine, it will have rapidly selected from its file the appropriate films for presentation. These films are already set up with a number of alternatives at each step, and with such inner workings that the machine is prepared to repeat, advance, or substitute material determined by the student's performance.

You will see from this one example that we are placing the machine and the subject matter in contact with the student, and vice versa, in a feedback relationship. Of course, we do not cover all possibilities; we do not even cover every possibility that a human teacher dealing with that one pupil could observe. But we handle a great many of the more common ones; we will strive for a very efficient and dynamically interesting presentation of a large amount of the material; we will do a very efficient job of examination of some of the student's understanding.

THE TEACHER'S ROLE

A brilliant student could romp through trigonometry in a very small fraction of the course time. A dull student would have to spend more time with the machines. The machines can be so set up that if a student fails to make progress at the required rate, he can be automatically dropped from the course. Of course, before that happens or before the brilliant student is allowed to complete the course, a special session with that student by a skilled teacher is indicated. But the teacher will be aided by having before him the complete records of what could be weeks of intensive machine operations.

This will make easier a personal study of that student's understanding and his way of thinking about the subject. The teacher will even be able to judge in what way the operation is inadequate and needs to be supplemented, both to take care of that particular student and to improve the automatic techniques. Some students will learn better than others with these machines. Ultimately, with the proper cooperation among experts in education, expert teachers, experts in trigonometry, and experts in engineering these automatic systems, we can evolve that high

level of match between the human teacher and the machine that we seek in that improved high school.

THE MEMORIZING MACHINE

We can further illustrate these concepts by other special cases. Let us take the memorizing machine, for example. It is important in many studies to do a certain amount of memorizing of facts and data. As a scientist, I know that a facility in study of an advanced subject oftentimes requires that background information be instantly available to the mind. But what a drudgery it is to memorize the weights of all of the chemical elements! In fact, about the best way to do this kind of memorizing is to get help from another individual who sits with the facts spread out before him and before whom the memorizer attempts to recite.

The memorizing machine could remove much of this drudgery and make it interesting and efficient. For instance, for the series of chemical elements, the machine could go through the list while the student punches out the corresponding atomic weights on a cash-register type of keyboard. When he misses one, not only does the red light go on (and the sign say "TILT!"), but the machine remembers that he has missed it. As it continues to chase through the list, it will throw in some of those questions that the student has already answered correctly, just to be sure and to give him the repetitive exercise, but it will more often come back to those where he had trouble previously.

A few minutes a day spent with memorizing machines, each of which is equipped with thousands of records to cover the important information to be memorized about various subjects, will probably accomplish more for the student than much more time spent in other ways.

Of course, it should be clear that this type of dynamic teaching and studying requires such a concentrated effort that it could not be used as the exclusive and total diet of the student, even if it had no limitations whatsoever. However, before we discuss these limitations, and before we try to make certain that we understand the fundamental difference that this kind of technical development could make in educational processes, let us take one or two other examples.

MACHINES IN THE LABORATORY

It is clear that the use of machines in which the student and the presentation are in responsive communication should be helpful in the presentation of theoretical concepts in science and mathematics, in the learning of basic principles, and the acquisition of information in most other fields as well. But what about such things as chemistry laboratory, English composition, and the teaching of languages?

Let us take the chemistry laboratory first—and remember that we are speaking here not of the principles of chemistry in the theoretical sense, which would be handled much as in the case of trigonometry, but rather of the physical handling of matter in the laboratory and the acquisition of appreciation of the scientific method of observation and deliberation. I think much can be done here.

TEACHING ON FILM

Picture, first, the student seated again before a special viewing screen and certain apparatus. The chemistry professor in the movie has the equivalent apparatus in front of him. He turns some valves and allows some fluid to go into a container. He adds to this another different fluid. He observes the characteristics of the combination, he refers to the theory, he describes what is happening and why it happens. He then asks the student to turn the valves in front of him to let so much red fluid into the glass below and so much blue fluid into the same glass. He tells the student that, if he has indeed poured the right amount in and observed every other requirement as described, he can expect certain results.

To show the possibilities, imagine that the instructor suggests that the liquid should be pink and asks the student to push Button A if he has obtained this result, and Button B if he has not. Now let us suppose that the student pushes Button A. The moment he does this, the film immediately switches to one in which the professor points an admonishing finger at the student and says, "Oh, oh, oh—now look at that liquid! That isn't pink. You were simply led by suggestion to expect the result. You didn't use your own powers of observation. Clearly, if you look honestly at that liquid, it is, if anything, slightly on the yellow side. You must learn the first principle of science. That is to be honest—not to expect a result, but to seek to observe what result you do indeed have and report it accurately."

On the other hand, if the student refuses to push Button A and pushes Button B instead, a different film will congratulate him on being objective and having the necessary characteristics for the scientific approach. It is quite possible for experts in chemistry and education, I believe, to create a large number of laboratory set-ups that can easily be kept full and ready by operators, so that the student can conduct his laboratory experiments without detailed supervision and with great efficiency and good records.

TEACHING ENGLISH

Teaching English and composition is difficult, as is instilling into one an appreciation of good literature. But even here we certainly can add to the exposure of the student to good literature and, by probing the student's understanding and response, we can alter the speed and nature of the presentation. We can improve his knowledge of the tools of good expression so that we leave only the more creative aspects (which must rest at least partially, presumably, on these tools and on the knowledge of the characteristics of good literature) to his personal contacts with the skilled teacher.

TEACHING FOREIGN LANGUAGES

Similarly, in the teaching of languages, vocabulary improvement, grammar, and understanding the spoken language could be advanced by these feedback machines. Even the ability of the student to speak the language could be enhanced by machine. He could respond orally to the animated film in front of him, repeat the foreign words spoken to him by an expert into the microphone, play back the results immediately, and repeat the whole process. This, you see, goes a substantial step beyond the use of records, which I understand is quite common in the courses now available commercially for learning foreign languages the "easy way."

Let us see what physical and human resources this high school would have. To begin with, the physical plant would include a large amount of apparatus that does not now exist but that can be designed and constructed with today's art.

There would be administrators and clerks who would handle all of the administrative processes but who would not be at all concerned with, and not be trained in, education. There would be a group of highly skilled teachers. The more conventional type of teaching would still be a substantial part of the total operation. For the new, automated material, these teachers would work closely with the experts on the subjects, and with the education engineers who design all of the electronic equipment that is basic to the process.

A NEW INDUSTRY

To back this up, of course, one would have a very substantial new industry in the United States concerned with the creating of these educational machines and the motion pictures and memory data used by the machines. In general, the industrial organizations concerned with the creation of machines that make possible the teaching of mathematics would have to employ experts in education, experts in mathematics, and experts in engineering. And this industrial team would have to be in good contact with the skilled teachers who make up the high-school staff in order that they might be able to improve their machines, create the proper material, and learn the shortcomings of all of their designs—either of the machine or of the material.

In addition, the high-school teaching staff would include education analysts, probably specializing in the various subjects. These individuals would go through the records of the individual students. They would be constantly seeking to discover the special problems that need special attention by the direct contact of teacher and pupil.

We notice a number of very significant points here. The high school becomes partially transformed into a center run by administrators and clerks, with a minimum of the routine assigned to the teaching staff. The teaching staff is elevated to a role that uses the highest intelligence and skills. A smaller number of teachers makes possible the education of a larger number of pupils. The creation of educational material moves partially out into industry, which goes into the education business in partnership with educators.

A NEW PROFESSION

There is probably a new profession known as "teaching engineer," that kind of engineering which is concerned with the educational process and with the design of the machines, as well as the design of the material.

One might imagine, for example, that a course in solid geometry, with its three-dimensional patterns, would be based around 3-D animated communicative-response systems and that some of the experts on the teaching of solid geometry should better be employed in industry than in the school. Those teachers who would be employed in the schools themselves would be individuals able to handle the more difficult problems that are left for the human teacher and the analysis of the processes that involve the use of the machine.

From the standpoint of the student, I do not know that his life need be changed in any fundamental way. It may be, of course, that the evenings and weekends would cease to be times for doing homework. The equivalent of homework, as well as the basic presentation periods, would be done perhaps during the normal working day, five days a week, with the evenings and the weekends used for the broader cultural, social, and athletic events. That is, the evening would be a time for a more relaxed participation in the learning and broadening programs.

I think it is true that, with this kind of an educational system, the student need not feel that he is dealing with cold machines in place of warm human teachers any more than he feels that way today when he reads a book by himself instead of listening to an oral presentation by a human teacher.

AN EXPENSIVE OPERATION

It is interesting now to look for a moment at the economics of such an educational system. In principle, it obviously has application to the lower grades and certainly to the university as well. But, wherever it is applied, it is quite easy to show that it is an enormously expensive operation.

Use that course in trigonometry, for example. As a motion picture, it would involve not one hour or two hours but, say, 100 hours of expensive teaching material. Unlike a motion picture, it would not be viewed by tens of millions of people all over the world. The audiences, for the most part, would be small, and unless we could unite all trigonometry students for a number of years in the common use of this same material, it is apparent that the cost would be rather large. If we pay something like 50 cents an hour to see an ordinary motion picture, then a trigonometry course would cost thousands of dollars per student, and the complete high-school year would cost tens of thousands.

SLAVES AND MACHINES

If we reflect for a moment on this matter of economics, we are reminded again that something is very wrong in the balance between that part of the national economy we devote to education and the part we are willing to devote to other things. To bring this point out, let us use the analogy of the Egyptians building the pyramids by the use of thousands of slaves pulling huge rocks—that is, by the most laborious and inefficient way possible in terms of the use of human beings. We imagine that we could have walked up to the Pharaoh and said, "This is not the way to build your huge monument. For this you should use bulldozers and cranes and steamshovels; why, a handful of men and a handful of machines would replace your thousands of human operators." But then the Pharaoh would have said, "Ah—but will it save me any money?" So we would figure it all out and discover that it would cost him much more if he changed to the machines because he paid his slaves so little.

In this system that I have described, we seek to elevate the teacher to the exclusive use of the higher abilities and qualities he possesses. It is a system that makes possible more education for more people with fewer skilled teachers being wasted in the more routine tasks that a machine should do for them. And we come up against this economic question. Today the teachers are doing all of these things—the routines and the handling of those levels of teaching requiring the highest of intelligence and training—and they are doing them for less than the cost of the machines, which could only hope to replace the lower level of the teaching art and skill!

A RIDICULOUS IDEA?

The examples I have presented here illustrate what I think is the most important point that can perhaps be contributed to a discussion of the relationship of the technological revolution and the educational process. While being defensive about these specific ideas, I won't go so far as to say that there is nothing in them. My work has accustomed me to the idea of being willing to allow imagination to roam freely, and my associates in science attribute a quotation to me which goes something like this: "Don't be ashamed to propose a ridiculous idea. Though worthless today, in 10 years it may be of no value whatsoever."

Technology and the Instructional Process

JAMES D. FINN

STATEMENT OF RATIONALE

THE STATEMENT of the rationale of this paper is best accomplished by presenting four propositions, as follows:

1. For purposes of analysis, the American educational system taken as a whole may be considered as a *society.*
2. The introduction, on a large scale, of technology into this (educational) society will be accomplished by the same types of stresses and strains, initiate the same types of problems, accomplish the same classes of objectives, and force similar reorganizations as the introduction of large-scale technology into any society, subject, or course, to modifications by the unique factors inherent in the specific (educational) society.
3. Because professional educators, for the most part, have not, up to this time, viewed current technological problems from this orientation, but rather (if at all) in specific contexts, *e.g.*, problems associated with the sudden introduction of television into education, it will be necessary not only to review thoroughly current educational theories from this more general point of view, but to create new theories encompassing and anticipating the new technological directions.
4. This new orientation toward technology and education will require rigorous analysis of the nature and possible roles of specific technological developments, the relation of these to educational purposes, and the prediction of effects on the society.

The remainder of this paper will attempt to delineate some of the elements of this technological study, to relate and analyze current and predicted technological developments, and to suggest some educational implications that might bear examination by the profession.

. . . .

THE POSSIBLE RELATIONS OF TECHNOLOGY WITH EDUCATION

Technology relates to education in at least three major ways. First, in a society in which science and technology are primary, such as in America, the society requires that the educational system insure an adequate supply of scientists and associated technicians. This requirement sets a curriculum problem, an organization problem, and many other problems associated with the screening, selection, and education of young people as potential additions to the nation's technical manpower.

Second, as a society becomes more and more technologically oriented and controlled, the question of the general education of all citizens is raised. The survival and management of the whole society theoretically requires more general education in the sciences and technology for all. Pressures arise for more mathematics and science to be taught to the entire population. Again, curriculum problems, organization problems, and a host of nagging, persistent general education problems arise.

Third, because of the tendency for technology to have no limits and constantly to extend into new areas, it is inevitable that, in an advanced technical society, technology should begin to extend into the instructional process itself. This is particularly true when, as will be shown, education has for a century or more been one of the areas of American society which has walled itself off from technological advances and, consequently, created a technological vacuum. That vacuum is now rapidly being filled.

This article is an abridged version of a paper presented at the special National Defense Education Act Seminar held by the Association for Supervision and Curriculum Development at the headquarters of the National Education Association, Washington, D. C., November 10-15, 1959. It was later printed in the Winter 1960 issue of the *Audio-Visual Communication Review* (Vol. 8, No. 1), "Automation and Education: III. Technology and the Instructional Process." The first part of the original paper contained a general discussion of the role of technology in society. The paper has been condensed by the editors of this volume to narrow its focus to the considerations that are directly relevant to the concern of the present volume with teaching machines and the programming of instruction. The author is professor of education, University of Southern California, and president (1960-61) of the Department of Audio-Visual Instruction, National Education Association.

It is within this third relationship—the application of technology to the educational, or, to be more precise, to the instructional process—that the balance of this paper is concerned. The three relationships just discussed—development of technicians, general education in technology, and the application of technology to the instructional process—cannot eventually be completely separated. However, the third relationship is sufficiently different to merit thorough analysis.

For purposes of this paper, a finer distinction will now be drawn. Within the educational process itself there are three general areas in which technology can be, or is being, applied. These are: (a) general administration, (b) testing, and (c) instruction. The uses of technical management systems, modern equipment, etc., represent the fairly obvious applications to the field of general administration. While it probably can easily be shown that this area of management is, taken as a whole, two or three decades behind its counterpart in industry, the problems associated with technology and school management are not as difficult as some of the others and will not at this time be considered. This is not to say that administration is not highly interrelated with the other two areas of testing and instruction, but the position can be taken that the problems arising from the latter two should guide the technical solutions in administration, not the other way around.

The second category, testing, is in many respects the most highly developed technology at present existing in American education. This is true both from a machine and from a systems standpoint. Further, the close relationships between psychological, achievement, and other types of testing and the instructional process are so well defined as to need no comment. However, with some exceptions, the technology of the instructional process can, for purposes of analysis, be isolated from testing. This arbitrary decision is made here in order to further the remainder of the discussion. We are left, then, with the instructional process by itself and can now turn to the impact of technology upon that process.

THE DEVELOPMENT OF INSTRUCTIONAL TECHNOLOGY

The development of a technology of the instructional process is relatively new. . . . In the pre-industrial phases of both education and industry, while industry was principally at the handwork, artisan level, the instructional process relied upon such devices as the slate, the hornbook, the blackboard, chalk, and limited single textbooks with few illustrations. Although, to keep this account of development brief, attention will be confined principally to some of the symbols of technology—equipment and machines—it should be emphasized that other factors—organization, etc.—were in the same state. As is well known, for example, the graded school is a late development.

At the beginning of the nineteenth century we note the famous changes in industry—the invention of a group of related machines for power, spinning, weaving, etc.—which made possible the factory system. Based on the work of Toynbee the elder, the term "Industrial Revolution" (now called the "First Industrial Revolution") is applied to this period.

However, in education the same revolution did *not* occur and instructional technology (with some exceptions here and there) remained at the pre-industrial level. During the last quarter of the century, there was some indication of a change. Oliver (1956) notes that, at the International Exposition in Vienna in 1873, "The exhibit of an American school with maps, charts, textbooks, and other equipment won admiration. The American school display at Paris in 1878 was even more outstanding." Significantly, Oliver's last mention of international attention to American instructional technology was at Melbourne in 1880.

By 1900, on the other hand, industry had established factories and was moving into assembly-line operations; had begun to apply, in a crude way, research and development concepts; had introduced the beginnings of modern management; and had developed a sophisticated financial system. Developments had reached such a state that the perceptive Henry Adams, the true prophet and seer of technology, foretold the coming of the Age of the Atom and the problems it would bring (H. Adams, 1918, 1919). In the years to 1950, aided by the acceleration of two wars, technology burgeoned and developed, piled machine upon machine, system upon system, added fantastically to power, and invented the method of invention. Technology transformed American society, philosophy, and art.

Technology, however, during the period from 1900 to 1950 only washed lightly upon the shores of instruction. During this same time span, when high-speed printing techniques, the radio, sound motion pictures, television, and other pieces of communication technology were invented, developed, and exploited, American education failed to apply these devices in quantity to the instructional process and failed to develop the appropriate technological systems necessary for this application. There were always rumblings, to be sure, as evidenced by the statement attributed to Edison in 1916 that the motion picture would replace the teacher. However, looked at from the vantage point of 1960, laboratories, project methods, libraries, and minute arrangements for audio-visual materials—the provisions to 1950—constituted what was still a pre-industrial technology for instruction.

By 1950 industry had entered what may be called the preautomation period which heralded the beginning of a movement toward true automation. At this time, much of the basic work on computers and other control mechanisms, including the mathematical and physical theories, had been done. At this time, also, mechanical systems of transporting work goods between machines had integrated some production lines into a system of mechanical or "Detroit automation." The period of automation was dawning, and I have arbitrarily set 1955 as the date at which this electronic-mechanical-systems analysis procedure began to be significantly applied. It is no accident that man began to push toward space at about the same time.

By 1950 American education had the potentiality, to carry on with the analogy, of a mass production technology. The hardware—projectors, recorders, television—and the materials were present. The systems concepts—saturation with audio-visual materials at the point of an instructional problem, for example, a concept derived from the military experience of World War II —were developed and known in a few audio-visual circles. A certain amount of incentive and public acceptance, also derived from World War II experience, could be drawn upon. The cake of custom, however, proved to be too tough, and the mass production stage, at least 100 years behind industry, was not entered except here and there on isolated little islands.

At this point, approximately 1955 . . . [the Ford Foundation gave] the instructional processes of American education a sharp push into a mass production technology. The time was ripe. There was a shortage of teachers; education and educationists were under fire from all sides; Neo-Technology was turning its attention to education; the educational race with Russia was under way; the natives were very restless indeed.

There are several interesting facets to this shove into mass instructional technology. First, television, both closed-circuit and broadcast, was chosen as the prime hardware. This was due, I think, to the unconsciously assumed basic concept of *mass production*, usually stated in terms of a shortage of teachers, large classes, and quality instruction. Second, the Ford impetus made provision, for the first time, for a technology of systems to go along with the hardware. Stoddard's own work in designing the school of tomorrow (1957) and the work of Trump (1959) and his associates in drawing the instructional images of the future for the American secondary school are essentially proposed systems.

Third, the Ford group made use of technology in creating the impetus in the first place. The technology of social psychology and public relations was drawn upon; educational decision-makers were the prime targets, and the means used, including high-priced public relations counsel, were the best available. This accounts for the fact that the teaching profession in general and audio-visual specialists in particular, really the only technologists of the profession, were, for the most part, left out of this move into technology.

. . . . . .

THE IMPACT OF PRESENT TRENDS

With this background of development, we can now come to grips with the impact of the present technology. Here our first principle should be recalled. If we consider education as a societal universe, what happens when we introduce energy into that universe suddenly and on a large scale? That is apparently what is happening. Two general effects, based on two analogous concepts, can be predicted. If the concept of "entropy" holds good in this connection, the educational system will become more highly organized and less random in nature (negative entropy will increase); if the observed historical-anthropological law of the introduction of a system of technology into a culture holds good, the distortions, stresses, and nonpredictable effects will increase until the culture becomes unrecognizable. Two examples—one taken from Mumford (1959), the introduction of the clock to regulate the habits of monks, and one from Muller (1957), the invention of writing in Egypt for the purpose of keeping accounts—changed the world several times over, and none of these effects were, of course, predictable.

What, then, are these present trends? If you don't count the continuous, grinding, sweating struggle carried on by audio-visual specialists, dealers, manufacturers, and producers to introduce, one item at a time, film libraries, projectors, recorders, etc., into the educational system in the slow process of conversion from a pre-industrial technology, two major trends can be identified. These two trends, at the moment, lead in opposite directions.

The first is the trend toward a mass instructional technology and is governed by machines and systems suitable for that purpose. Foremost, of course, is television, of which there are four instructional types: (a) broadcast on an educational channel, (b) broadcast on a commercial channel, (c) closed-circuit of the Hagerstown-Penn State type in which live instructors are used either to supplement instruction or to provide direct instruction exclusive of classroom teachers, and (d) the Compton type in which filmed lectures are distributed via the closed-circuit medium as replacement for classroom teachers. In all cases, the desire is to reach more students with fewer teachers, or "quality" in instruction.

Mass instruction technology in another form includes, of course, the massed film systems. The prime example is the EBF [Encyclopedia Britannica Films] series in physics and chemistry, amounting to over 300 half-hour motion pictures intended to be used where there are no science teachers, where there are teachers not considered qualified, as audio-visual aids to more qualified teachers, and with very large groups of students. These films, of course, may theoretically be used over television. The exploitation of the overhead transparency projector as a device for mass instruction in English grammar and composition in the experiment at Newton, Massachusetts, is another example of mass instructional technology.

In opposition to this trend of mass instruction is a growing technology for individual instruction. This trend is the audio-visual wave of the future (for the moment). The most dramatic development here is that class of systems and instruments known as teaching machines. Actually, I would class all teaching equipment designed for individual or near-individual operation as being in the category of teaching machines. At present, then, there are approximately five types, listed here on an ascending scale of sophistication: (a) individual reading pacers and similar devices; (b) individual viewing and listening equipment for existing slides, filmstrips, motion pictures, and recordings; (c) language laboratories of all types; (d) specifically programmed printed verbal materials such as Crowder's (1959, 1960a) scrambled textbooks, the programmed textbook (Glaser, Homme and Evans, 1959), etc.; and (e) true teaching machines of the Skinner or Pressey type containing carefully worked out verbal or pictorial programs with various ingenious mechanical or electronic arrangements to test student reaction, inform him of his progress, errors, etc.

EXTRAPOLATION I

These two major trends toward technologies of mass and of individual instruction are not science fiction describing the future; they are with us now. Assessment of this impact is our most pressing and difficult problem. It doesn't take much of a crystal ball to see that a combination of these two technologies is the next immediate step. For example, let us take Maurice Mitchell's new series for his chemistry course on 150-plus films. Let us further assume that Mr. Mitchell makes a deal with a manufacturing company to program their multiple-choice Didak teaching machine with material to go along with the chemistry films. . . . Let us continue our assumption by saying we will project the films hourly on the Compton College closed-circuit television film distribution system and place the machines in proper cubbyholes for Compton College students to work at specified times. We could get, almost immediately, to what can be called total educational automation.

Or, in another context, let us suppose that the new Ford creation, the Learning Resources Institute, initiates another Continental Classroom in mathematics over national broadcast television; and that the new audio-visual teaching machines now being developed by Hughes Aircraft are programmed to relate to this new series; and that these machines are placed by the Foundation in every public library in America; and that, at the conclusion of the instructional period, students report to designated centers where the Educational Testing Service will examine them and certify them to the colleges of their choice—or high schools, if you will—as having passed the course. Think, for a moment, about that one. It is now possible not only to eliminate the teacher, but the school system.

These may be considered extreme applications, but such applications, at least experimentally, are inevitable. In fact, one such study, testing the combination of these two technologies, is now under way. The point is that both the mass instruction systems and the technology of individual instruction—teaching machines—are gathering terrific momentum. These technologies are going to hit education with a million-pound thrust. What will be the effect upon our educational society? What is the role of the teacher, the audio-visual specialist, the curriculum director? *Quo vadis* the curriculum itself? This last question is particularly pertinent if the medium or image governs the message, as McLuhan (1959) maintains.

EXTRAPOLATION II

Technology never sits still. The extrapolations given in the last section are immediate. What of those that can be made for a time slightly farther into the future? I see several developments. First, it is reasonable to expect that teaching machine development will move from verbal-Socratic-Skinner type programming to audio-visual-branching type programming. That is, the machines will present, based upon student pretests, conceptual content using films, slides, filmstrips, tapes, and/or video tape as the medium [cf. Ramo, 1957]. The presentation sequences will be longer and the student will be given an opportunity to select additional sequences for further explanation if the machine, through testing, informs him that he needs it. Records, of course, will be maintained instantaneously by miniaturized computers.

Secondly, the work in England of Ashby, Pask, and others with teaching machines based on the design concepts of biological computers will affect the technology of the present teaching machine. . . . Pask has already produced a machine which, in training key punch operators, actually senses the characteristics of the student as he works and automatically adjusts the program of the machine to the individual needs of the student. This is a pure transaction, and Pask (1958a) maintains he can develop such a machine to teach decision making with present hardware and know-how.

A second development of the future begins next year with the stratovision experiment of the Ford Foundation over a city in Indiana. News of this experiment is gradually filtering into educational circles and has appeared here and there in the press. One or more airplanes, carrying multiple television transmitters circling overhead, can receive multiple signals from the ground and broadcast them over much greater distances than conventional tower transmission. The programs will go into several states and, because of the location, into both urban and rural areas and into innumerable heretofore independent school districts. Effects of this on local control of curricula, curriculum-planning procedures, images of master teachers, patterns of school organization, etc., have yet to be speculated about, let alone explored. Nothing technical would prevent, by the way, a national network of such aircraft as an addition to our other national means of transmission.

A third predicted development is that it is technically possible, I have been informed, although not yet economically feasible, to produce a motion picture film with characteristics similar to the Polaroid Land still camera film—instantaneously developable. What this would do to video-tape, to production techniques, and to the possibilities of local production of materials is of great interest to the audio-visual field.

The fourth development that can be anticipated is the most spectacular. Essentially, it is based on a systems concept. It is theoretically possible now to design an automatic classroom under the control of the teacher. Most of the elements are present or can be designed. Such a classroom would have total light and air control, automatic projection and television systems, technical provision for the best possible discussion environments, display situations, etc., which could be changed at will. By planned programming the classroom could be made to function for major presentations, small group discussions, individual work at teaching machines, creative periods, etc. All of this *could be* under the control of the teacher. The classroom then would become the teaching machine. Adrian TerLouw, at Eastman Kodak, has done some work in this area and it is, of course, implied in the provocative article by Simon Ramo (1957) which is so often cited in discussions of teaching machines. An acquaintance of mine refers to this concept as the "mad scientist" classroom.

If we look at this series of possibilities, we can sense a change in trend. Of the four cited, only one—the stratovision experiment—is clearly on the side of mass production. The other three lean toward the technology of individual instruction. The automatic classroom is a combination of both, but one in which the human element still plays the central part with the machines being the slave of man, not the other way around.

IMPLICATIONS

To assess the implications of this current and predicted impact of technology on the instructional process is not easy. As with any such speculation, it is also hazardous. It is, however, a job we must attempt.

We have to start, I think, with the proposition that, based upon the historical development of instructional technology, our educational society is in the position of a backward or underdeveloped culture suddently assailed by the twentieth-century engineer. In addition, reactions in some professional circles to the advent of television are similar to those factory workers of the nineteenth century who attempted to destroy the machines that were replacing their jobs. In the long run, the machines remained; instructional technology is, no doubt, here to stay. Our problem becomes one, not so much how to live with it on some kind of featherbedding basis, but how to control it so that the proper objectives of education may be served and the human being remain central in the process.

Certain things are obvious. I think the concept of negative entropy will hold. The thrust and energy of technology will force a greater organization upon us *at every point it is applied to instruction*. Such an arrangement as the stratovision experiment will require a tighter and different organization in those five or six states and in those many school districts of the Midwest. The closed-circuit experiment in Anaheim, California, just beginning, is already, it seems to me, making requirements for organization—and scope and sequence in the curriculum, of the district itself, and of the nature of the television presentations—the rigor of which has never been felt before in Anaheim.

Programming, in its machine, computer, automation sense, is a matter of extreme organization—the piling up of energy. Most of the devices and systems of the new technology require, one way or another, this type of programming—television, teaching machines, mass films, language laboratories. This means that programmers are needed. Who is to do this? This is essentially, in the old sense, both a curriculum and an audio-visual problem. It is also a social problem. The heartland is programming. He who controls the programming heartland controls the educational system. Will it be a foundation, a committee of scientists, textbook publishers and film producers, the NEA, the school superintendent, the board of education, the students, or the general public?

Too, how many of us will go overboard and sink with the old concepts that will be absorbed or outmoded and tossed to the sharks by the new technology? Take the concept of instructional materials which some curriculum specialists love so well and which caused quite a controversy in DAVI [Department of Audio-Visual Instruction, NEA] a few years ago. The fight was futile, as is the love. The concept of programming and the systems and systems analysis it implies completely absorb the idea of materials. "Instructional materials" becomes an outmoded atomistic, pretechnological concept useful mainly to the historians of education. (McLuhan [1959] maintains the same thing about audio-visual aids.)

The concept of audio-visual education may accompany instructional materials down the drain, or it may not, depending on whether or not it can be redefined acceptably.[1] The Skinner- and Pressey-type teaching machines and their descendants are, for example, primarily verbal devices, and yet their management, programming, etc., as technical electronic devices belongs somewhere. At a practical level, within a school situation, someone is going to have to worry about them.

It is my position that the audio-visual field is in the easiest position to help integrate these mechanisms properly into the instructional process. They are not primarily audio-visual; they are primarily technological. The audio-visual field, I think must now suddenly grow up. The audio-visual specialists, are, of all educational personnel, the closest to technology now; we have, I think, to become specialists in *learning technology*—and that's how I would redefine audio-visual education.

---

[1] Redefinition is the subject of a paper now at press.

What of the curriculum people? Their work is just beginning. It can become infinitely more rigorous and, at the same time, more satisfying. I suspect some of the emotional Rousseauists will have to get out of the curriculum business. The rest can step in and face this problem of programming a learning technology in human terms and find great excitement and reward.

. . . .

THE PHILOSOPHY OF ADVENTURE

In closing, I should like to return to the position I took in my Los Angeles address to the DAVI national convention in 1955. Whitehead has said that it is the business of the future to be dangerous. Technology certainly has made this aphorism into the outstanding fact of our time. Technology is now making the future of instruction capricious and hazardous. But, in doing so, it has presented us with more opportunity and more choices than ever before. If the future is an adventure, it is an adventure *because* of technology. The cost of civilization is the fact that we can make wrong choices because of the alternatives technology presents. The reward of civilization is the freedom provided by technology and the opportunity to make the right choices. This cost and this reward we now face with the technology of the instructional process. We *must* look forward to the adventure and not present what Herbert Muller (1957) noted as "the curious spectacle of civilized man forever marching with his face turned backward—as no doubt the caveman looked back to the good old days when men were free to roam instead of being stuck in a damn hole in the ground." The "good old days" are gone forever; approached with intelligence and zest, the days of the future will be better.

PART V

Some Recent Work

During the past two or three years a large number of individuals throughout the country have been stimulated to undertake active work concerned with teaching machines and programmed learning. The primary impetus for this rapid expansion of activity has unquestionably been the writing of Skinner. Though some of this work was initiated shortly after the publication of Skinner's 1954 paper, there were at first few published articles to indicate the widespread interest that had been generated. During 1958, however, numerous papers began to appear, and the trend has continued to increase during 1959 and 1960. The articles included here in Part V represent a sample of papers which have appeared recently, in addition to those by Skinner and his associates at Harvard presented in Part III.

Some of these additional papers represent preliminary thinking and investigation, and about half of them have not been formally published heretofore. Most of the papers have been presented at professional meetings or submitted to sponsoring agencies as reports of ongoing work. The recent work has been of several kinds. First, a number of individuals have written articles primarily devoted to the exposition of basic notions expressed in the earlier papers by Pressey and Skinner. Some of these have been particularly concerned with educational implications, and some have stressed underlying rationale and relation to theory. Other individuals have been primarily concerned with developing and trying out instructional programs, either on a small-scale experimental basis or, in a few instances, with the aim of programming entire course units for immediate practical use in school and college classes. Still others have presented analytic discussions of important variables in the programming of material or in the design of teaching machines, and a few have reported controlled experiments to investigate the operation of various factors or procedures in programmed learning.

The first three papers in Part V all report ongoing work in the programming of college subject matter. The first paper, by J. W. Blyth, is based on an extensive project being undertaken at Hamilton College. In addition to his specific experiences in the programming of a course in logic which the author describes, this paper provides a very readable general account of machine-presented instruction. One special point of interest in Blyth's paper is the author's cogent discussion of emotional reactions to the "mechanization" of education, which some fear the use of machines might bring about. The advantages of machine-presented programmed material are concretely illustrated by the author's interesting account of his own experiences in using them in the classroom. Blyth also presents suggestions for the arrangement and packaging of program units and other aspects of practical implementation.

The second paper, by J. A. Barlow, was provided at the request of the Editors as an informal, interim report on the comprehensive programming effort currently in progress at Earlham College. It was felt that such a report would be useful in conveying some of the working details that enter into a large-scale attempt to put teaching machine concepts into practice in ongoing instruction at the college level. An interesting brief report of an intensive effort to program part of a single course in engineering is given by Mayhew and Johnson in the third paper. This paper is of special interest in indicating the considerable labor involved in converting a standard course into a machine program.

An exploratory attempt to program a brief sequence in arithmetic concepts for use in the fifth and sixth grades is reported by Keislar in the next paper. Of interest are the individual records of performance shown for several subjects. This study highlights the necessity for careful analysis of student responses to the steps in a program. In this case, as the author points out, many of the steps were too difficult for many of the students. The necessity for early revision of a program on the basis of student performance is thus emphasized as a basic rule in program development.

The paper by Glaser, Homme, and Evans reports one of the recent studies on programming conducted at the University of Pittsburgh. The paper describes and illustrates the "programmed textbook" developed as a simplified device for implementing self-instructional procedures and studying programming variables without the use of "hardware." The "programmed textbook" and the standard textbook are compared as aids to teaching, and characteristics and variables that enter into the construction and investigation of programmed learning sequences are described.

The next three papers report experimental efforts designed to investigate properties of programmed learning sequences. The paper by Evans, Glaser, and Homme describes a pilot investigation in which the size of the "steps" was varied in a programmed learning sequence, and in which overt (written) and non-overt modes of responding were compared. The performance of students who used the programmed text was also compared with that of students who used standard text material.

Coulson and Silberman report a detailed experiment in which the three factors of response mode (multiple-choice vs. constructed response), size of step, and use of branching were each manipulated as experimental variables. Two conditions for size of step were employed. The "small step" condition utilized the first 104 items of the program described in the paper by Holland (p. 215-28); for the "large step" program, about half of the items in this program were deleted. The branching program similarly deleted certain items from the program, but only when the student had already responded correctly to related items. The experimenters employed a simulated teaching machine (cf. Rothkopf, 1958a, p. 327) in which an experimenter concealed behind a screen manipulates the program sequence so as to obviate the use of a complex machine. It is interesting to note that the effects of the variables studied were generally more marked in their influence on instructional time than on the level of achievement attained.

The study by Fry provides another comparison between constructed and multiple-choice responding. It should be noted that the materials to be learned by Fry's subjects consisted of rote vocabulary items presented as paired associates, rather than a sequenced program of meaningfully organized material as in two previous studies. Fry's study, unlike Coulson and Silberman's, yielded significant differences favoring the use of constructed responses. These resulted in better terminal test performance (on a constructed-response criterion test). Fry introduced three methods of controlling for the difference in time expenditure required by the two modes of response. Constructed response showed superior performance under all three conditions. The interaction between achievement and instructional time required, in this and the two preceding studies, suggests the use of a measure of learning effectiveness which would take into account both the amount learned and the time expended—perhaps a ratio of amount learned per unit time.

The paper by Gilbert is one which will certainly invite controversy, both because of its content and because of the author's rather flippant style. It has been included here, even though some of the more specific notions presented are quite debatable, because of the emphasis given by the author to two important concepts: First, that teaching-machine instrumentation should be governed by prior discovery of what is really needed in order to teach a subject matter—rather than biasing the teaching method by the characteristics of preselected machines. Second, that the requirements for successful programming of a subject matter can best be discovered by observing the responses that learners make during trial attempts to teach them. In his trenchant style, Gilbert proceeds to give an amusingly presented, though seriously intended, set of "rules" for accomplishing the job of programming.

In the latter part of the paper, Gilbert suggests a more radical approach for the discovery of laws of behavior through the utilization of a so-called "culture-free" laboratory, in which extensive, long-term observations of behavior would be made. Although many features of this proposal are highly speculative and controversial, the serious point is nevertheless made that traditional laboratory experimentation in psychology may often have biased or predetermined the results it obtained because of the way in which variables were selected for study. He contrasts this "traditional" situation with the notion of a laboratory conceived to reveal the operation of unanticipated relationships and variables that are important for the effective control of learning. Gilbert has also written several earlier, unpublished papers in which concepts and procedures for the construction of programmed material are stated in concrete, operational terms. Some of the main points in several of these papers, which have been rather widely read and have influenced current concepts of programming, are summarized in Appendix I.

The paper by Homme and Glaser continues the emphasis on procedures for program construction. The authors are impressed by the dearth of explicit rules for a technology of programming, and suggest a preliminary classification of program elements which might serve to facilitate the development of a specific technique for program preparation in various subject matters. They present a variant of Skinner's (1958c) illustrative program sequence in high-school physics, rewritten according to the system they suggest. The authors also comment on the problem of analyzing student behavior prior to the preparation of a programmed learning sequence.

The paper by Amsel reacts to some of the ideas set forth by Skinner. Amsel discusses, in terms of theoretical concepts, the notion that the difficulty level of steps in a program should be kept low so that the occurrence of incorrect responses during learning is minimized. He examines the role of errors in the acquisition of knowledge, and presents an analysis of two circumstances in which incorrect responses might be treated differently depending on specific conditions. The second major point concerns the scheduling of reinforcement in programmed learning. Amsel examines the relevance of experimental work on reinforcement schedules, and distinguishes between the kind of schedules that appear suitable for meeting different objectives appropriate to early and later stages of training.

Pressey's recent paper represents his reaction to the current trend of development in programmed learning. Parts of it will evoke marked disagreement among many individuals currently active in work on programming and in the study of individual self-instruction. Much of Pressey's concern centers around what he interprets as too great a preoccupation with the use of "small step" programs for all educational purposes. He also calls into question the comparative advantages of constructed as opposed to selective forms of response. Such problems, Pressey would feel, need to be attacked by well-planned and vigorous experimentation, using alternate approaches toward the implementing of basic features of teaching-machine instruction. In addition, he stresses the need to prepare the student for subsequent learning in coping with a nonprogrammed real world—a concern which is certainly shared by many of those whose approach to the use of teaching machines differs from that of Pressey. In the main, the reservations which Pressey expresses concerning the approach and philosophy of current workers in programmed learning do not appear to reflect a disagreement over the basic advantages of using self-instructional devices as a means toward more efficient learning. He shares with more recent exponents of teaching machines a questioning, experimentally-oriented attitude toward conventional teaching practices. However, he also emphasizes the importance of trying to identify and build on such gains as have been made by educational innovators in improving various aspects of education during the past few decades.

The last two papers in Part V attempt to summarize some of the important factors that underlie the effectiveness of teaching-machine instruction. The paper by Lumsdaine identifies and tabulates some major classes of variables, and analyzes similarities and differences among a number of teaching machines with respect to such factors as mode of response, type of response comparison, nature of feedback, program variation and display characteristics. It then discusses in more detail the question of size-of-step, techniques of prompting or cueing in automated learning programs, and some of the research problems these pose. The final paper, by Carr, also surveys the characteristics of programs and devices. He provides a detailed review under a series of carefully organized headings, including device-display or input characteristics, response or output characteristics, characteristics of the confirming mechanism, reinforcement characteristics, program characteristics, and principles and sources of error in program preparation.

Teaching Machines And Human Beings

JOHN W. BLYTH

I. INTRODUCTION

HUMAN aspirations and ideals are intimately related to human emotions. The emotional attitudes of approval and disapproval in a given culture are usually expressed through certain words. It is often difficult to disentangle the attitudes expressed by those words from their descriptive meaning. Since the Industrial Revolution a continuing complaint has been heard that the use of machines has led to the debasement and the materialism of our civilization. The word "machine" has thereby acquired a heavy emotional burden. The same burden is now rapidly piling up on words such as "automation." Thus the phrase "automation of education with teaching machines" represents such a summation of horrors for some people that it blocks intelligent inquiry into the merits of teaching machines.

I believe this accounts for the paradox that some reporters, editors, radio and television commentators, whose very jobs were created by such developments, utilize the most modern automated machinery to disseminate their uncritical emotional reactions against machines. Fortunately, the mechanization of the media of communication has not mechanized the minds of all those who have something to communicate. No more will the mechanization of the method of presenting instructional materials mechanize the materials to be presented. In the end, the widespread use of teaching machines will not depend on emotional reactions to a word but on the machines' value in helping us to cope with urgent educational problems.

This paper was prepared in connection with a project at Hamilton College, under sponsorship of the Fund for the Advancement of Education. It is an expanded version of a paper given in October 1959 at the annual meeting of the American Council on Education and has also been published in the April 1960 issue of *The Educational Record*. Dr. Blyth is professor of philosophy at Hamilton College, where he is collaborating on this project with Charles Godcharles, of the Department of Psychology, and others.

A teaching machine is simply a mechanical device for presenting to a student a succession of instructional items requiring some discriminative response and providing the student with an immediate check on the accuracy of his response. Mechanical tutors of this kind can free human tutors for other teaching tasks wherever instructional materials can be satisfactorily programmed for this method of presentation. This has obvious implications for any country either unable or unwilling to reward its teachers well enough to recruit them in adequate numbers. On this score alone, the potentialities of the teaching machine are worth examining and exploiting as rapidly as possible if they prove their value.

There is another less obvious but no less important advantage to be gained in exploiting the teaching machine. The machine makes it possible to provide some of the conditions that we have long known to be necessary for efficient learning. A savage instructing his son in the use of the bow and arrow knew that the son needed plenty of practice and that he had to see where the arrow went on each shot if he was to make any improvement on the next. The importance of this immediate feedback was formulated years ago by Thorndike as the "law of effect." According to this law of learning, an action which leads to a satisfactory result tends to be repeated. In the contemporary terminology of B. F. Skinner, of Harvard University, immediate reinforcement or reward is important in the learning process. As far as the application of this principle is concerned, we do a much better job of teaching rats, pigeons, and football players than we do of teaching mathematicians and physicists. Teaching machines designed for individual use make it possible to provide this immediate reinforcement for every student.

Any improvement in the control of conditions providing more efficient learning should have important consequences. It may be possible for a given instructional unit to be learned more thoroughly and easily in less time. To state the same point in another way, a student should be able to learn more material more easily and more thoroughly in a given unit of time. The further development and rapid exploitation of the teaching machine is one of the most promising approaches to the problems of teaching the increasingly complicated body of modern knowledge to ever larger masses of people. The teaching machine offers hope that this may be done without abandoning our democratic ideals of education and lowering our ambitions to the education of an intellectual elite.

II. TEACHING RATS AND HUMANS

Full exploitation of the potentialities of the teaching machine will require a great amount of work by many people. The number of research projects related to the teaching machine is rapidly increasing. The kind of research project most likely to produce quantifiable and convincing results is one relatively small in scope and concerned with a limited number of controllable variables. Much more of this is needed. But not all of the important questions can be stated in terms suitable for a tidy laboratory research project. If we look only for problems which can be so formulated, certain of the most important issues may be ignored. Furthermore, some attempts to use the teaching machine in regular courses will be helpful in determining some of the research problems likely to prove most valuable.

Laboratory experiments have demonstrated that the use of immediate reinforcement techniques is a very efficient way of training animals to make a discriminative response to selected stimuli. Rats and pigeons can be trained to perform a whole series of such responses in sequence. Their actions then form a complicated "chain" of events.

In certain areas in human learning a drill technique has traditionally been employed. The mechanical presentation of the questions and answers has an obvious application wherever it is felt appropriate to use drill methods, but it is not so obvious that it has any wider application.

Psychologists do not offer a curriculum of courses in maze running and bar pressing. There is, to the best of my knowledge, no science of maze running to be taught. The order of right and left turns in a maze is arbitrary for each maze. The only reason a rat should turn to the right rather than the left at a certain point is that it is that turn which leads to reinforcement. No better reason can be learned because there is none.

Students sometimes pass courses in logic and mathematics in the same way. They learn to write down certain words or numbers in answer to questions because those are the answers which produce reinforcement in the form of passing grades. Frequently they cannot imagine why the teacher prefers these answers to other equally attractive answers, but they hope to keep the teacher happy by giving the answers they have learned he wants.

A student who gives a particular answer merely because that is the one the teacher reinforces has not learned a subject such as mathematics or logic. I make no claim that knowing mathematics and logic requires some special reasoning faculty or involves some mysterious insight into a special world of numbers and logical entities. But the reason for saying "valid," for example, must be more than the fact that saying "valid" caused the teacher to produce a reward. Reinforcing a response is part of the cause for the student's learning. But if the student has learned only that the response causes reinforcement, he has not learned logic or mathematics. The answers the teacher will reinforce are not selected by arbitrary decision or whim. They are the right answers, in the sense that they are required by the rules of the relevant system. To learn mathematics and logic, the student must learn to produce answers because they are the right answers. A familiar difficulty for the teacher is to discriminate right answers given for the right reason from right answers reached by guesswork. Only the former should be reinforced.

Mathematics and logic are two of the most highly organized and integrated subjects that confront students. Both require much more than rote memory. Thus logic constitutes a good test for determining the usefulness of teaching machines in developing complex high-level intellectual skills.

Last year, with the help of a grant from the Fund for the Advancement of Education, John Jacobson and the writer started an experiment in teaching logic at Hamilton College. Our experience has convinced us that the most effective teaching aid now available is a program of questions and answers designed for a teaching machine and that the most efficient method of presenting such a program is in a machine using microfilm and designed for individual use.

The use of the programmed materials we developed gave us a new experience in teaching. We found many advantages and no disadvantages. Among the advantages we noticed the following:

1. We wasted no class time on routine checking or on drill.
2. We wasted no class time on unprepared students. We knew in advance exactly who had done the work on the programs and who had not. We believe that evidence of work on programmed material should become a ticket of admission to class.
3. Examination of student work on the programs enabled us to prepare for a class period with prior knowledge of the points that needed further clarification.
4. In the classroom work we could presuppose a common background of experience with a large number of relevant examples encountered in the programs.
5. We wasted no time trying to locate and correct misconceptions students had acquired through unchecked practice.
6. We could usually count on a working command of basic concepts and principles. Class time could therefore be devoted to further development of the concepts and their application to new areas.
7. Classroom efficiency was increased by at least one-third. Instead of a standard three hours a week we met two hours a week. Yet we were able to cover more material more thoroughly than ever before.
8. Individual differences were not entirely removed, but the range was greatly reduced, with every student moving closer to mastery. I had at least three students in a section of 20 who, in my judgment, would have failed the course in the preceding year. They passed with a very safe margin.
9. Fewer private conferences with individual students were needed. In the conferences that were necessary it was possible to diagnose the difficulty very quickly by analyzing the student's answers to the programmed materials. More often than not, such conferences led to revisions in the program in the hope of avoiding such troubles for other students.
10. The programs constitute an excellent diagnostic instrument in locating individual and class differences; an analysis of student answers provides an equally good

diagnostic instrument in locating shortcomings in a program. After testing and revising programs a certain number of times, we should have programs which we know from experience will enable the vast majority of students to master complex materials in a minimum time.

11. There was a great increase in interest and improvement in morale.
12. One student taught himself the first semester's work with little help from us.
13. During the summer of 1959 some of the materials prepared for Hamilton College students were used in a program of advanced studies for gifted junior and senior high-school students presented by the Board of Cooperative Services in Oneida County, New York, in another project aided by the Fund for the Advancement of Education. Examination results compared favorably with those of college students.

III. PROGRAM REQUIREMENTS

Probably every teacher has at some time been baffled because a student could not see a point that was perfectly obvious to the teacher. When this happens, a good teacher knows that it is necessary to back up and to lead the student step by step until he can see the point for himself. In its simplest terms this procedure indicates what is required for a good program. Two traditional techniques are combined in modified form: the Socratic method of teaching by asking questions and the Cartesian method of analyzing a problem into its smallest parts and proceeding from the simple to the complex.

However, the kind of analysis needed goes beyond the Cartesian analysis of a problem into its smallest logical parts, which is necessary but not sufficient. In symbolic logic one can find some of the tidiest, neatest systems ever developed. In presenting a system, we may start with the simplest elements, a few rules of operation, a few axioms and postulates, and then deduce everything in rigorous order. Yet a course in symbolic logic taught in this tidy fashion is probably one of the most baffling courses many students will ever encounter.

There are indeed some situations in which it may be best to start with relatively simple elements and then put them together to form more complex units. For example, in propositional logic we may start with a simple statement such as "the window is open." We may then construct larger units by connecting two simple statements with words like "and," "or," and "if . . . then." It may then be shown that the truth of the larger unit depends on, or is a function of, the truth of its parts.

But if all problems were treated in this manner, disaster would result. Suppose we try to teach the parts of an argument in this fashion. We present the statement "the door is open" as an example of a conclusion. Then we present "the window is not open" as a premise. Next, we present "the door is open, or the window is open" as a premise. This would probably confuse the student at first, but with practice he could learn to call it a premise. Finally, we are ready to put these elements together to form a complex unit called an argument. We then have:

The door is open, or the window is open.
The window is not open.
Hence the door is open.

If he has mastered the preceding steps, the student will confidently identify "the door is open" as the conclusion. We must now baffle him by saying that in some cases "the door is open" is a premise rather than a conclusion.

The trouble here, of course, lies in the fact that the statements constituting parts of an argument are not classifiable out of context. The same statement may be a conclusion in one argument and a premise in another. Furthermore, there is no single clue which can be used as a reliable guide even in context. The conclusion is sometimes stated first, sometimes in the middle, and sometimes at the end of an argument. It is sometimes in the middle, and sometimes at the end of an argument. It is sometimes identified by the words "therefore," "hence," or "thus." Sometimes it is not.

In this illustration we simply cannot start with isolated simple parts. We must start with a whole argument and train the student to recognize the parts as they appear in context. To do this we can start with some examples in which the required discriminations are easily made. Initially the discrimination may be made quite obvious by some device such as underlining or otherwise setting off one part and asking questions requiring direct attention to that part. Further examples lead through major variations. Extra clues are reduced, and the cases gradually become less obvious. After a student has worked through a program of such carefully selected and graduated examples, his perceptual habits are modified. He no longer sees merely a set of statements. He begins to see premises and conclusions as related parts of a larger complex unit.

This illustration points up one of the most important advantages of a teaching machine. In the textbook an author is necessarily restricted to the use of a very few examples to explain a point. Unfortunately, every example exhibits many different characteristics, and it is not always easy for the student to distinguish the relevant from the irrelevant features of the illustrations. The easiest way to learn to make the necessary discriminations is to examine a wide variety of examples. In a very short time a well-organized sequence of items on a teaching machine can provide the experience needed for the learner to identify the relevant characteristics for himself. Furthermore, in such a sequence of questions it is possible to start with examples familiar to every student and to lead him step by step to develop general principles and surprisingly complicated formulae for himself. Since he has derived the general principles from his own experience with specific cases, there is no problem of sheer verbalism or the mechanical manipulation of meaningless symbols.

In working on textbook exercises, a student who has trouble with one problem is frequently tempted to skip on to the next. More often than not, this leads to increasing difficulties. The teaching machine presents the items in a controlled sequence of graduated steps. The student always has the information needed for the next step. If a particular sequence of questions does not, in fact, develop the intended understanding, then it should be revised until it does.

I am not suggesting that the use of textbooks and classrooms is obsolete. However, it is important to note that the questions in a good program are not drill questions based on prior textbook readings or lecture notes. The subject matter is learned through answering the questions. The numerous examples encountered in the programmed material provide the background experience necessary for ready comprehension of a text or a lecture, for intelligent participation in discussions, and for the application of what has been learned to new areas.

The considerations mentioned above suggest the following requirements for a program of questions prepared for machine presentation. The program must start with something familiar to the student. With machines using microfilm, it is possible to provide some of the necessary background experience through such means as pictures, charts, diagrams, and maps. There must be examples of sufficient variety if the essential features are to be distinguished from irrelevant details. The questions must be graduated in difficulty so that the student may take each new step for himself.

There is much discussion over other features of a program. Some urge the use of multiple-choice questions. Others recommend questions requiring a student to construct an answer. For certain purposes multiple-choice questions may be quite satisfactory. On his first trip through a maze, a rat must make a guess at each fork in the path. There is no better way to select a path, for the connection is completely arbitrary. In human learning there are also many arbitrary connections to be learned. It is possible that multiple-choice questions are satisfactory in such contexts.

However, not only mathematics but all subjects taught in school have some structure imposed by the subject rather than by the whim of a teacher. Where structure is important, the guessing encouraged by multiple-choice questions is surely the wrong way to arrive at a right answer. For this reason we have been using constructed answers almost exclusively in our logic program. The wording of the questions and the context set by the sequence of questions are usually adequate to indicate the type of answer required. Occasionally this is not so. Then the quickest way to set the context is to indicate the relevant alternatives.

Some machines are designed to rerun a set of questions a second time—presenting only those questions missed on the first run. This process is continued until all questions have been answered correctly. Again, this method may be efficient where the connection between the question and answer is wholly arbitrary and must be mastered by drill.

However, there are two reasons for not following this routine in programs for structured subject matter. In the first place, the context is very important in a carefully graduated sequence of questions. If a student misses Question 16 because he has not gotten the point of the 15 questions leading up to it, he can only be more baffled by getting the same question repeated out of context. He needs to go through all the steps again or, better yet, through an alternative sequence using a different approach. In the second place, in such subjects as logic and mathematics nothing is of less importance than memorizing that "valid" or "394" are the answers to certain specific questions. What is important is to learn that "valid" is the correct answer under certain conditions. A mistake on Question 16 is of no consequence if the student profits from it and successfully handles subsequent problems similar to that of Question 16.

Programming techinques may vary in other ways as well. One variation is illustrated in the scrambled textbook. This technique requires multiple-choice questions. Each alternative directs the student to turn to a certain page to see whether or not his answer is correct. On the pages for the wrong answers he may find explanations of his error, further clues to the right answer, and perhaps a little scolding. The page for the right answer will contain congratulations and directions to the next question.

This technique is intriguing in some respects, but it is prodigal in the use of time and materials. For example, a sequence of 40 questions with only two alternative answers per question would require 40 extra pages in a text or frames in a film to carry the wrong answers. This appears insignificant in connection with an isolated experimental set of questions. However, its economic importance becomes clear when one must think in terms of a possible 10,000 questions for one course and thousands of copies of a program prepared for widespread use. To justify the large extra material costs, the advocates of this technique must demonstrate a very marked superiority over a straightforward program. Furthermore, the advantages gained by explaining the reason for an error may be obtained by presenting it along with the correct answer. This is indeed helpful in situations where common-sense misconceptions lead most students to the same wrong answer.

Another proposed technique involves a complicated internal structure for the program with corresponding complications in the machinery. The questions presented to the student are in part contingent on his prior responses. Suppose, for example, that a set of 10 questions is inserted in a program to check a student's mastery of a certain principle. If he runs into any trouble, presumably he needs more work on the same principle. A computer with a complex memory circuit and indexing circuit can be programmed to present further examples of the same principle. However, if the student runs through the key set of questions with no difficulty, it is presumed that he has mastered that unit. The computer may then offer the student a choice among several branches in the program. One branch may lead to the next unit. Other branches may involve an extension of the concepts just mastered or various applications of those concepts.

In dealing with the problem of individual differences, this branching technique has certain obvious advantages. Again, any widespread use of machinery for automatic branching is economically unfeasible. However, the advantages of this technique may well be obtained in another way. Practical considerations suggest packaging a program in relatively small lesson units. A quick check of the results on one such unit may be used to indicate which lesson is to be offered next.

After having considered the practical requirements for widespread use of this, the characteristics indicated by the psychology of learning, and the structure of materials to be learned, we would suggest the following general system:

1. Programs should be packaged in relatively small units.
2. Cost of duplication is important. Film is easily reproduced at low cost.
3. The packages must be easy to store and to use without being damaged. This may be done by inserting strips of film between plastic sheets.
4. Machines must be easy to operate.
5. A group of machines may be arranged in a room with a check-out desk. A student arrives at his scheduled time, picks up the appropriate lesson, inserts it in the machine, and goes through the lesson. He then returns the lesson to the desk.

6. At the same time that the student picks up his lesson, he obtains and signs a score card, which is inserted in the machine for automatic scoring. This score card has several functions: (a) It gives the student information about his work on that lesson. (b) The teacher will find it helpful in diagnosing individual difficulties. (c) A glance at the cards for the whole class will locate difficulties requiring classroom attention. (d) Records easily derived from these cards will be very useful in evaluating and revising sets of questions.
7. All programs to be used in the visual machine would be packaged in a standardized manner. The same machine would therefore be suitable for any subject.
8. An audio unit may be plugged in for automatic coordination of visual and auditory instruction. This will permit the construction of programs using any combination of visual or auditory questions requiring either written or oral replies. This has obvious advantages for instruction in language or music.

IV. IMPLICATIONS FOR OUR EDUCATIONAL SYSTEM

It cannot yet be said that the effectiveness of the teaching machine has been fully demonstrated. However, if it turns out to be as valuable as early evidence indicates, there are many possible consequences for our educational system. Suggesting a few of these will serve to emphasize the importance of further investigations.

Even a modest increase in efficiency of the learning process has very significant cumulative effects. For example, if we express a 12.5 percent gain in terms of time, the average student would master in 14 years what he now learns in 16 years.

Any gain on the time dimension will be accompanied by a gain in mastery. By present standards it is a gentlemanly achievement to do approximately 70 percent of the work in a course satisfactorily. A grade of 60 in one course is usually considered adequate evidence of preparedness for the next course. Difficulties for student and teacher are cumulative when the 30 to 40 percent of unlearned material is needed for work at the next level. Satisfactory programs presented by the machine should permit every student to master the essential elements at one level before proceeding to the next.

Correlative with a gain in mastery is a decrease in failure. Instead of eliminating scholastic failure by automatic promotion, we may be able to eliminate it by nearly automatic achievement. Furthermore, students who discover the pleasure of success in school are unlikely to feel compelled to develop skill in throwing rocks through school windows.

It should be possible to develop programs for teaching machines which would produce a genuine mastery of basic skills at a very early age. It is by no means inconceivable, for instance, that high-level mathematical skills should become commonplace rather than the result of specialized training for a few.

The individual teacher should welcome this form of teaching aid. He can cover more ground more effectively in fewer hours spent in the classroom and conference room. Furthermore, the time he does spend with students individually or in groups may be used much more effectively. He knows just what can be assumed as a basis for further work.

An increase in general efficiency has a double significance for both private and public institutions. The reduction of class contact hours for a course from three hours to two hours a week represents a saving of teacher time and a saving of classroom space. Both factors are becoming more and more important as the school population and the costs of education simultaneously increase.

Inasmuch as the teaching machines discussed here are designed for individual use, they do not need to be installed in large rooms or even in school buildings. It might even turn out in the long run to be more economical to place a teaching machine in every home than to double the size of school buildings.

Finally, some possible consequences for human values should be mentioned. It should be kept in mind that the teaching machine is merely an instrument for presenting instructional materials. In effect, it provides superior private tutoring for every individual. The tutoring will be superior because the programs presented will be the work of experts who have tested the programs on many students and revised them until they do the job intended.

Many people fear that the use of machines will somehow mechanize the people using them. This is as absurd as supposing that a person is being turned into a machine by listening to a hi-fidelity stereophonic recording of the New York Philharmonic Orchestra. It is the teaching machine that operates when a button is pressed. It is the student that presses the button. In 15 minutes' work on a good sequence of questions and answers the average student probably devotes more concentrated, alert, and active attention to the instructional material than he does in the course of an hour in the average class. The correct answer is immediately available when he needs it, and the next question is presented as soon as he is ready for it. In a classroom situation it is impossible to obtain optimal timing for all students at once. Thus there is a prospect that the use of good programs in teaching machines may develop more alertness and spontaneity than the average classroom situation.

Our democratic goals of education are frequently criticized for having brought about a leveling down of standards and achievement. Problems created by increasing numbers have led some persons to abandon the democratic goals and to concentrate on training an intellectual elite. The large-scale use of tested programs in teaching machines may make it possible to cope with the growing problems and, at the same time, raise our standards and actual achievements nearer the point of mastery.

Recently I heard a woman complain that her son had been assigned to a teacher no one liked, and she was afraid her son would not work well for this teacher. This is a phase frequently forgotten in praising the inspirational qualities of a good teacher. Good programs presented by machines will surely not hamper an inspiring teacher; but they should diminish the number of such comments as, "I would have been able to learn physics if I had only had a good teacher when I started mathematics."

Finally, it is worth noting that a program presented by a machine has a dual kind of objectivity. In the first place, every student answer is either right or wrong, or in some respect either satisfactory or unsatisfactory. This is not at all the same thing as a true-false test in which the student must respond with the words "true" or "false." If a student is asked to identify the author of a sonnet, the answer is either right or wrong. In order to progress, the learner must know whether his performance is correct or incorrect. A sequence of questions and answers which provides such knowledge at each step guarantees objectivity.

In a good program of questions and answers the student should have little difficulty constructing the correct answers. The difficulty lies in the construction of the questions and organizing them in sequence. The teacher must learn to ask questions which elicit answers that both teacher and student can easily recognize as correct or incorrect. If a teacher cannot ask any questions requiring the student to make some discrimination that is clearly right or wrong, it is doubtful whether that teacher is teaching anything but shifting whims.

Programs presented by machines offer another kind of objectivity. Failing students frequently rationalize their difficulties by charging that the teacher doesn't like them. A machine develops no likes and dislikes. It has equal patience for the slow and the quick. It does not discriminate between rich and poor. It makes no distinctions with regard to race, color, or creed. We do not yet know what part of our educational program should be presented in this manner. But it is clear that thoroughly tested programs will make it possible to come closer to realizing the democratice ideal of equal educational opportunity for all.

The Earlham College Student Self-Instructional Project: A First Quarterly Report

JOHN A. BARLOW

OCTOBER 1958—EARLHAM COLLEGE

B. F. SKINNER, of Harvard University, gave a public address on "Teaching Machines." Three or four members of the faculty had heard of the work in this area prior to this time, and their prediction that it would be a stimulating and controversial topic proved correct. In the evening about a dozen faculty members gathered for coffee and conversation. Few were convinced of both the feasibility and the desirability of wide use of teaching machines; all were intrigued by the possibilities and impressed by Fred Skinner.

JULY 1959

With a grant from the U. S. Office of Education and with the stated intention of raising additional funds from nonfederal sources, Earlham College started a three-year research project entitled "New Instruction Media: Self-Instruction, Guided Instruction and the Role of the Teacher." Thus Earlham joined Hamilton College, and the two became the first colleges to undertake major research projects in automated instruction on a widely interdepartmental basis.

The sequence of events between these two dates is a function of the ethos of a specific campus and the interests of specific individuals. However, primary general factors would appear to be the small liberal arts college's emphasis on undergraduate teaching and the constant interaction between members of different departments. By the time of submission of the formal proposal in February to the U. S. Office of Education, 14 individuals in different academic departments had drawn up preliminary statements of some of the possible applications of teaching-machine principles to their teaching fields.

The preliminary plans for the project may be summarized under the following headings.

PROBLEM

To explore procedures for improving the effectiveness of the teacher:

1. By using mechanical and electronic aids for guided self-instruction
2. By programming individual and group course work in the light of modern learning theory
3. By studying the interrelations of the various kinds of teaching procedures in order to determine the best use of each.

OBJECTIVES

To shed further light on the following questions as general questions and in connection with specific areas of instruction, courses, and portions of courses:

1. How can college instruction programs in widely varying disciplines most profitably incorporate recent advances in learning theory?
2. How can the learning sequence in various areas be best programmed so as to make efficient use of instructional devices operated by the individual student?
3. To what extent can these devices be adapted so as to provide increased individualization instead of increased uniformity in the programming of instruction?

PROCEDURE

1. Analysis of present educational procedure in specific areas in the light of modern learning theory

---

This is a slightly edited report of work in progress, as of October 1959, on a project at Earlham College supported by a grant from the U.S. Department of Health, Education, and Welfare, Office of Education. Dr. Barlow is chairman of the Department of Psychology at Earlham and is coordinator of this project.

2. Construction or arrangement of new programs and of appropriate devices for the instructional process which will provide optimum feedback to the student and instructor
3. Establishment of a self-instruction center
4. Initiation of specific course and area experiments using Earlham students as subjects
5. Re-evaluation of the role of the teacher and of the student-teacher relationship.

OCTOBER 1959

Three months have passed since the official confirmation of the government grant and the start of the project. The fruits of the initial stages which are immediately apparent are:

1. Project personnel, aside from the central figures of the subject-matter specialists involved in programming, consist of coordinator, technical associate, clerical assistant, shop technician, and student assistants.
2. A report on a preliminary experiment conducted during the spring in the genetics section of general biology indicated that a simplified form of the Pressey immediate feedback multiple-choice test did not significantly affect quiz scores over this small unit, but the type of testing met with considerable student approval.
3. One month during the summer was spent on preliminary planning and programming by instructors in introduction to music and elementary Russian.
4. Released time this first fall has been provided for programming in freshman English, genetics, elementary Russian, and statistics.
5. In elementary Russian and in statistics the instructors are programming the course work and using it immediately on current classes. In psychology, the Holland-Skinner program is being used. In general biology, genetics, and introduction to music certain topics are being presented in programmed form.

6. Released time is definitely scheduled for Spanish in the spring of 1960 and for introductory chemistry starting in the fall of 1960.
7. Through the shop at Indiana Medical School Institute of Psychiatric Research, and with the help of Charles Ferster, we have been supporting some exploration of the possibilities for an inexpensive and easily portable "semi-automatic tutor."
8. A 30-booth project audio laboratory for research and instruction primarily in languages was in operation by the beginning of the 1959 fall term. This laboratory at present involves primarily the best current features for more or less conventional instruction. Modification of the procedures for use of this equipment following the new "teaching machine" principle will precede and accompany modification of the equipment itself as proves necessary or desirable as the project develops.
9. Offices for the project and small self-instruction center to house experimental teaching machines and to provide a research area for this year are being constructed in one of the college classroom buildings.

At such an early stage in the project all procedures are tentative. Certain preliminary general procedures appear to be worth noting as a result of their preliminary success.

We have kept our equipment expenditures limited to items which we are sure will be used. As construction of our programs progresses, additional equipment required to handle *these programs* will be purchased or designed.

In order to introduce the subject-matter specialists who are to do the actual programming to what is already known, a file of pertinent material is kept available and as up to date as possible. Each programmer so far has himself worked through at least a portion of the Holland-Skinner program for the natural science psychology course at Harvard. The programmers thus learn some of the background of the basic principles we are currently attempting to follow at the same time that they become familiar with the oldest program available.

Rather than attempting to make our preliminary devices for presenting the programmed material cheatproof or so constructed that the student is forced to do as the programmer wishes, as a result of the nature of the self-instruction device provided, we have explored the possibilities of teaching the student the desired behavior by a brief training sequence. We have found it helpful to devote one class period to an "experiment" on learning with and without knowledge of results. Half of the class are blindfolded and practice drawing three-inch lines for 10 trials without knowledge of results; the other half of the class practice drawing the lines under the same conditions but with information as to how long each line actually is immediately after the trial. (See Experiment 1 in Tinker, M.A., and Russell, W. A., *Introduction to Methods in Experimental Psychology.* New York: Appleton-Century-Crofts, 1958.) The students seem to react quite favorably to assignments involving sequences of question-answer-confirmation after such an introduction and demonstration.

We look forward to many possible changes in the over-all picture of our college program as our own research, and that being done at other schools, progress.

We look forward to the day when all students will be required to pass vigorous proficiency examinations in various basic areas before admittance to advanced courses. By the college's providing programs in basic skills, all students may be required to have or to acquire a basic command of the material recommended for college preparation (for instance: English language skills, foreign language, basic mathematics, general science concepts, and even such skills as typing). Proficiency in these skills would be prerequisites to entrance in appropriate college courses, and work in building up these skills need not be given any college credit. As a result of inadequate background, some students may need to spend one or more summer sessions working on such basic programs; but the college instructor of the future may be able to depend on every one of his students having certain required skills to a designated level of minimal proficiency.

At all levels, degree of objective proficiency rather than grades and credit hours may be recorded in areas of work in which this turns out to be feasible. Instead of a certain number of credit-hours with a specified minimum grade in each (and thus a tremendous spread in background), a certain specified terminal level of accomplishment may be required as a prerequisite for advanced work in many areas.

If the work done on self-instruction devices is regarded as "independent study" and the student is oriented toward his own responsibility to complete this work, a transition to greater emphasis on other types of independent work in all areas of study may be evolved. We expect our use of self-instruction through mechanical and semimechanical devices to be accompanied by a deliberate attempt to encourage self-instruction in other ways. It is a matter of college policy to develop an over-all program which will involve a period of off-campus study for most students in the junior or senior year (this is already in practice in a strong and rapidly developing foreign study program in connection with our present curriculum). As the climate of learning changes from the more conventional lock-step which is involved necessarily in much of our current program, we anticipate a total re-evaluation of our entire campus life—curricular, co-curricular, and extracurricular. These possibilities are so far-reaching that they are still merely loose speculations.

Teaching Machines

D. J. MAYHEW and A. F. JOHNSON

IN THE Mechanical Engineering Department of the University of Utah, a pilot study is being made of the use of teaching machines in engineering education. B. F. Skinner's article in *Science*, October 24, 1958, is the direct source of our interest.

The essential characteristic of the teaching machine, as the term is used here, is its ability to locate and respond to the particular instructional needs of the individual student using the device. The machine aspect of the device consists in the functions of:

1. Routing a variable sequence of instruction statements and questions to the student, and
2. Keeping accurate and conveniently tabulated records of the responses of individuals and groups.

The program which the machine handles is set up for the following sequence:

1. A graph, equation, or statement of instruction is presented to the student.
2. He is asked a question requiring interpretation of this material. (Student writes answer on paper strip integral to machine.)
3. The strip with his answer is advanced under a window to prevent alteration, and the correct answer appears at another window.
4. The student pushes either a "correct" or "incorrect" button telling the machine which branch of the program to continue on and recording the student's ability or inability. The machine "remembers" wrong answers and subsequently introduces extra instruction and questions in these areas until right answers result.

The similarity of this sequence to what goes on in a human teacher-student relationship recommends our serious consideration of its use in this age of educational ferment. Some potential advantages of its use appear to be:

1. Reduction of faculty lecture time, permitting increased faculty-student counselling time.
2. Reduced educational cost per instruction unit.
3. Freedom from present course-scheduling problems.
4. Students proceed at their own pace.
5. Accurate up-to-date records of each student's comprehension are readily available.
6. Capacity for steady evolutionary improvement of the machine program by tabulating the responses of each student and thereby indicating weak places in the program.

Perhaps the machine can never have the "human quality" necessary if it were to replace the human teacher. It is too early to consider complete replacement. To those who react with aversion to the words "machine teacher," it would be better to consider the device as simply a very talented textbook. We have not attained sufficient understanding of the educational process to reject such new ideas on theoretical grounds. We must set up controlled experiments. Such experiments assume an ability to measure the results of an educational activity. Many people question that we have any such ability. But we do give grades, and largely on the basis of impersonal test scores. It seems we have already assumed some ability to measure. The subjection of the learning process to scientific experimental techniques has long been resisted because of its complexity. Our educational crisis shows a necessity for improvement. Necessity is famous for her children.

Our preliminary study of machine teaching included spending 60 hours to develop a trial sequence of instruction, questions, and answers as a machine program to give the content of *one* of our present lecture + text assignments on column failure. It is easy to calculate the "energy barrier" to be overcome before a standard course can be converted to machine programs. The users of electronic computers have set us a very good example here. Their programs also require much time and skill to develop. They have established program libraries and organized cooperation in developing programs of interest to many users.

This article is reprinted from the *Journal of Engineering Education*, Vol. 50, 1959. The authors are members of the faculty of the Department of Mechanical Engineering, University of Utah.

The scope of our undertaking requires the combined efforts of many engineering educators and coordination by a national committee of the ASEE.

The Development of Understanding In Arithmetic by a Teaching Machine

EVAN R. KEISLAR

THE USE OF teaching machines for the teaching of spelling and arithmetic combinations has already been shown to have merit (Skinner, 1954; Pressey, 1927). And studies have demonstrated that automated teaching can result in more than simple rote learning (Porter, 1957; Ferster and Sapon, 1958). In this study the problem was to explore the possibility of using a multiple-choice method for the automated teaching of "understanding," specifically, an understanding of areas of rectangles. By *understanding* is meant the ability to answer a variety of questions different from those encountered during training but belonging to the same general class; the broader this class is, the greater is the understanding.

APPARATUS

The teaching machine used in this study was an extensive adaptation of the Film Rater used by the Navy for teaching aircraft identification. Multiple-choice items on a Kodachrome strip-film were projected in sequence upon a viewing plate. The learner responded to each item by pressing one of five buttons. If the answer was correct, a green light was turned on and the next item could be brought into view by pressing a special button. But if this answer was wrong, a red light came on; only after turning off this red light could the learner try again. To proceed to the next item, the learner had to answer correctly.

A special device recorded a graph of all right and wrong answers for each item. If the wrong answer was given, the pen moved to the right one-twentieth of an inch. For each correct answer the pen moved vertically an equal distance to a new line. Hence the subject's performance for any item could be read from

---

This paper originally appeared in the *Journal of Educational Psychology*, Vol. 50, 1959. Dr. Keislar is with the School of Education at the University of California at Los Angeles.

the horizontal line of the graph corresponding to this item number.

THE ITEM SET

The total program consisted of 120 items, 10 of which instructed the learner how to operate the machine and informed him of the goal to be attained. The remaining 110 items were constructed to provide a sequence beginning with concepts of squares, rectangles, length, and width. Following items requiring the pupil to indicate the number of square units in rectangles, the concept of area was presented. Applications included paint coverage, rug size, and tile-laying, followed by practical problems of adding and subtracting areas and finding the length or width of rectangles. The set concluded with items involving cost.

PRINCIPLES OF PROGRAMMING

Several of the principles outlined by Skinner (1958c) and illustrated in his completion-item set were applied to this multiple-choice approach. To illustrate the principles discussed below, a short sequence from the program, Items 18 through 29, is presented in Figure I.

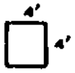
1. The step from each item to the next in the sequence should be small enough so that the learner almost always gets each item right. Although Homme and Glaser (1959b) and Coulson and Silberman (1959b) found that smaller steps resulted in better learning and took less time per step than larger steps, definitive evidence on this issue has not yet been obtained. In this study it was assumed that if a pupil selected the wrong alternative to an item, he did so either because he was improperly selected for the program or because of inadequate prior learning in the program itself. The programmer should make sure that a pupil learns enough before an item is presented so that generalization to this item is practically assured.

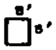
In opposition to this line of reasoning is the argument in favor of higher item difficulty that the learner is encouraged to "formulate his own hypotheses and try them out." This procedure may have merit if information is supplied to the student showing why the alternative is incorrect (Crowder, 1959) or if branching in the program permits special remedial instruction. But where, as in this study, the pupil is informed only that he is wrong when he makes an error, he is given no more information than when he gets the item correct. One fifth-grade subject, after completing the program in this study, commented: "It's hard to know why you get something wrong. When I got it right, I knew it. When I got it wrong, I didn't know why." Since the absence of an explanation is likely to heighten the aversive consequences of failure, it appeared most desirable to adopt a minimum difficulty level for the items in this study. If it seemed helpful to emphasize that certain responses were wrong; instead of having the student learn this by "being wrong," special items were constructed for this purpose, e.g., "Which of the following figures is NOT a rectangle?"

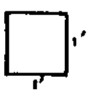
2. Skinner's use of the vanishing stimulus and the prompt facilitates the occurrence of the correct response. For example, in Items 25 and 26 colored rows of squares encourage the right answer. In Item 27 these become merely dotted lines which vanish completely in Item 28. In Item 20 the correct answer "1 ft." is prompted by the dimension "1 ft." in the diagram.

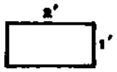
3. To promote generalization within a broad class of items, a process necessary for understanding, the learner should acquire a variety of verbal responses which might later be used, through intraverbal associations, to evoke other appropriate responses. This is simply a process of meditated or secondary generalization, a complex example of which is found in Judd's theory of transfer through verbal principles. For instance, it was judged that if the pupil learned to group squares within a rectangle by rows, as in Items 25 through 28, he would acquire intraverbal associations which would promote the correct response in learning later in the program to multiply the length by the width to find the area and, still later, to divide the area by the length to find the width. In other words, with these items the pupil is being prepared to "gain insight into why" you multiply in one case and divide in the other. Appropriate intraverbal associations, such as verbal principles, definitions, or characteristics, should function to extend the pupil's learning to entirely new items which involve the same principle or concept. If this indeed can be accomplished, the use of multiple-choice items in

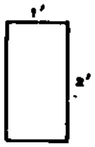
18. Squares can be large or small. But if each side on a square is 1 foot long, this is called a square foot. Which figure is one square foot?

A. 

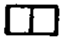
B. 

C. 

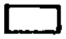
D. 

E. 

19. Here are two squares: They are exactly the same size. Suppose I put the squares together side by side, like this:



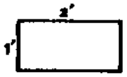
What kind of figure will the two squares together make if I leave out the middle line like this?



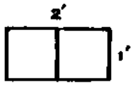
A. Another bigger square

B. A rectangle

20. Here is a rectangle. It is 2 feet long and 1 foot wide.

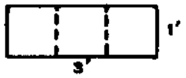


Suppose I draw a line in the middle, like this: I will make 2 squares. How long will each side of each square be?



A. 1 ft. C. 3 ft.
B. 2 ft.

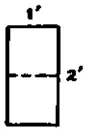
21. Suppose you have a rectangle which is 3 feet long and 1 foot wide, like this:



How many squares, 1 foot long on each side, can I make out of this rectangle?

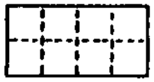
A. 1 C. 3
B. 2

22. How many square feet are in a rectangle which is 2 feet high and 1 foot wide, like this:



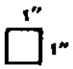
A. 1 C. 3
B. 2 D. 6

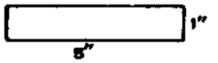
23. How many square feet are in a rectangle which is 2 feet wide and 4 feet long?



A. 2 C. 6
B. 4 D. 8

24. Each side of this square is 1 inch long. This is called 1 square inch. How many square inches are in this rectangle which is 1 inch wide and 5 inches long?





A. 1 D. 4
B. 2 E. 5
C. 3

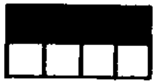
FIG. 1. A sample sequence of items, Nos. 18-29, to illustrate programming principles.

(cont. on p. 429)

automated teaching results in something more than "mere recognition" of the right answer.

4. Procedures or concepts which are not otherwise involved in the sequence of items should be reviewed periodically. For example, a review of the process of grouping squares into rows, originally presented in Items 25-28, was provided later by six items occurring at intervals throughout the program.

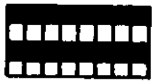
25. Here is a rectangle which is 2 inches wide and 4 inches long. How many square inches are there in this rectangle?



There are two rows with 4 squares in each row. How can we write the total number of square inches?

A. 2+2 C. 2+4+2+4
B. 2+4 D. 4+4.

26. This rectangle is 4 inches wide and 7 inches long. How many square inches are in this rectangle?

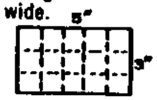


A. There are 4 rows with 4 squares in each row.

B. There are 8 rows with 4 squares in each row.

C. There are 4 rows with 7 squares in each row.

27. This rectangle is 5 inches long and 3 inches wide.



There are 3 rows of squares. In each row there are 5 square inches. How many square inches are there in the rectangle?

A. 5+5
B. 5+5+5
C. 5+5+5+5.

28. Here is a rectangle which is 5 inches wide and 9 inches long. How many square inches are there in the rectangle?

A. 9+9+9+9+9
B. 5+5+5+5+5
C. 5+5+5+5.
D. 5+5+9+9+9.

29. A rectangle which is 5 inches wide and 9 inches long contains 9+9+9+9+9 square inches. What is a quick way of finding how many square inches this is?

A. Add 9 and 5
B. Add 9 and 9 and 9 and 9
C. Multiply 9 by 5
D. Divide 9 by 5.

FIG. 1 (Cont.). A sample sequence of items, Nos. 18-29, to illustrate programming principles.

5. Other techniques included the repetition of the correct answer on the succeeding item (as in Item 29), the irregular appearance of "interesting" colored pictures accompanying the item, and the use of a variety of forms of the multiple-choice item.

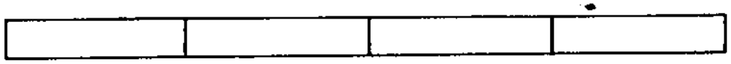
SUBJECTS

Fourteen experimental subjects and 14 controls, individually matched on the basis of intelligence, sex, reading ability, and pretest scores, were selected from the fifth and low sixth grades. All subjects showed competence in multiplication and division but little acquaintance with the topic of area. A fifteenth subject who had completed the program was dropped from the study because of an automobile accident prior to the post-test. The control subjects were given no special instruction of any kind. They were used to control for the effects of incidental learning such as that which might result from the administration of the pretest.

PRETEST AND POST-TEST

Both these tests were of the free-answer or essay type. The pretest consisted of 12 problems involving multiplication and division, in addition to eight problems dealing with areas of rectangles. The post-test consisted of the same eight problems of the pretest on area plus another eight, most of which were more difficult. Sample problems on the post-test were:

1. A gallon of paint will cover an area of 200 square feet. How long a stretch of fence could you paint with 2 gallons if the fence is 5 feet high?
2. A sheet of cardboard is 3 feet long. It weighs exactly 4 ounces. What is the area of this cardboard if it is 2 feet wide?
3. A sheet of paper is 10 inches wide and 12 inches long. This sheet of paper is cut into strips. The strips are laid end to end and then joined with Scotch tape. This long strip now looks like this:



Can anyone tell what the area of this long strip is? ____________. If so, what is it?________________.

PROCEDURE

Experimental subjects operated the machine for two or three periods on successive days. The total time spent with the machine ranged from one hour and 30 minutes to slightly over two hours. The post-test was given the day following the end of such machine instruction to each experimental subject and his control.

RESULTS

The mean score of the experimental group on the post-test was 12.4 with a standard deviation of 5.6. The corresponding control group mean was 5.4 with a standard deviation of 3.7.

Since the post-test scores for the experimental group showed a much greater variance than did the control, a sign test was used. All except one of the experimental subjects showed a higher post-test score than did their matched controls, a difference which is significant at the .01 level. Although the experimental group answered every item, except one, of the post-test better than the control group, most of these pupils missed several items which were similar to those presented in the program; they learned far less than what had been expected.

The total errors on the program ranged from three to 118 with a mean of 54. Jane's record, presented in Figure II, was the poorest performance and indicates very little learning as measured by the post-test. Part of her problem may have been inadequate ability in computation since her computation score was relatively low. Even for a typical student like Kenneth, whose record is shown in Figure III, the program was too difficult; 40 mistakes is entirely too many according to the criterion adopted. The program appeared to be ideal for Byron, whose record appears in Figure IV. Although he showed little ability in the field of area on the pretest, Byron's performance on the post-test was outstanding. He was able to generalize from his training so well that on the post-test he solved completely new problems of obtaining the area of a parallelogram and a triangle.

The rank order correlation between total number of errors on the program and the gain on the post-test was – .83. While this, of course, supports the hypothesis that the optimum difficulty level of each item should be low, it does not permit, in itself, any such conclusion. The rank order correlation of mental age was .52 with gain on the post-test and – .79 with number of program errors. On the basis of this limited sample

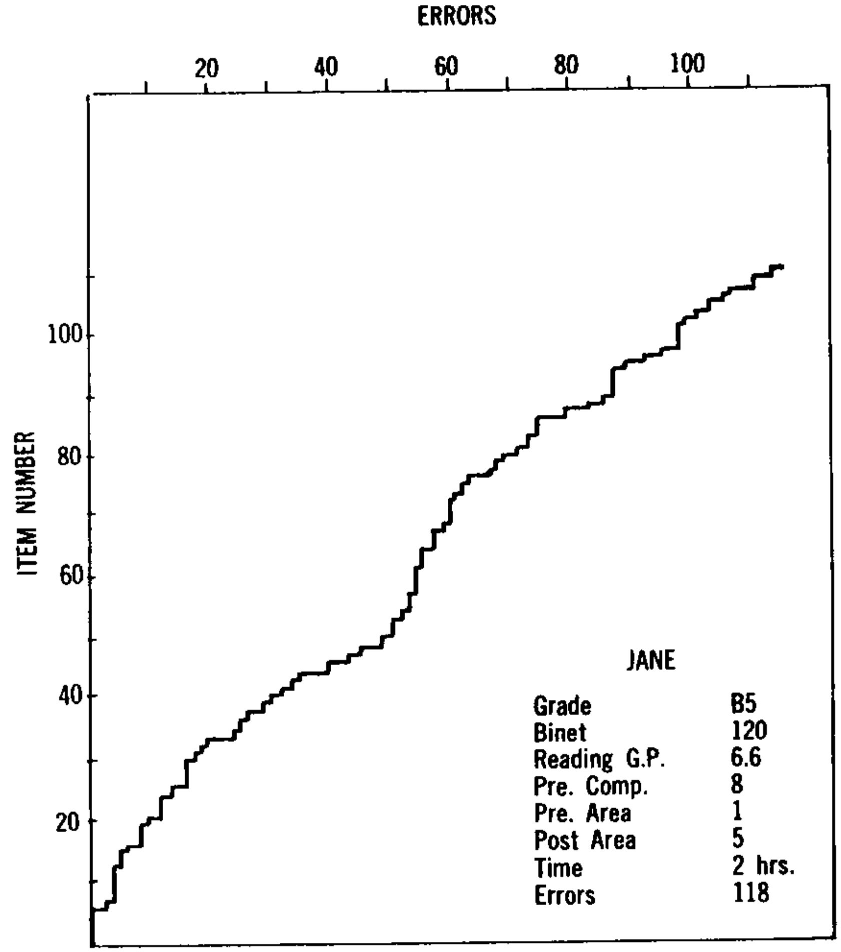


FIG. II. Graphic record of the poorest performance

it appears that the program was more appropriate for the brighter children.

DESIRABLE REVISIONS IN THE PROGRAM

In the absence of definitive evidence on the question, it appears that the major weakness of this program was that it was too difficult for most of the pupils. Revisions should be made along the following lines:

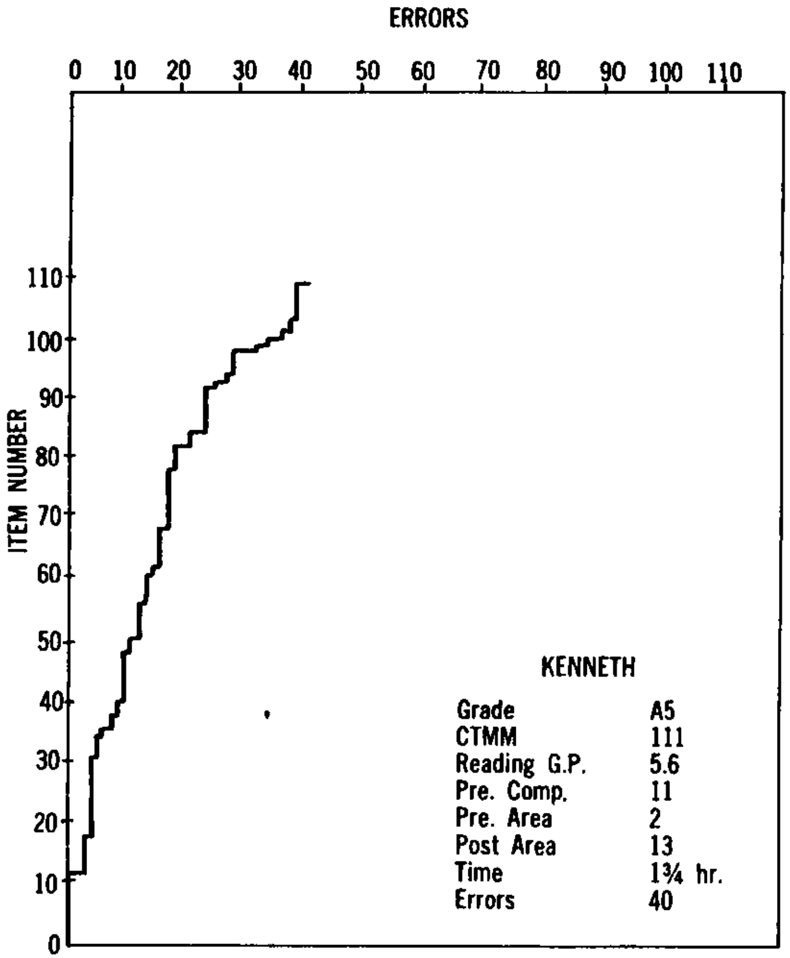


FIG. III. Graphic record of a typical pupil

1. Since the reading load was probably a major obstacle for many pupils, sentences should be shorter and the total amount of reading less for each item. Possibly several versions of the program, at different reading levels, would be desirable.
2. The steps in many, if not most, cases could be smaller. For example, Item 22, which was missed by three pupils, could be preceded by items analogous to 19 and 20. Item 24, also missed by three pupils, could be divided

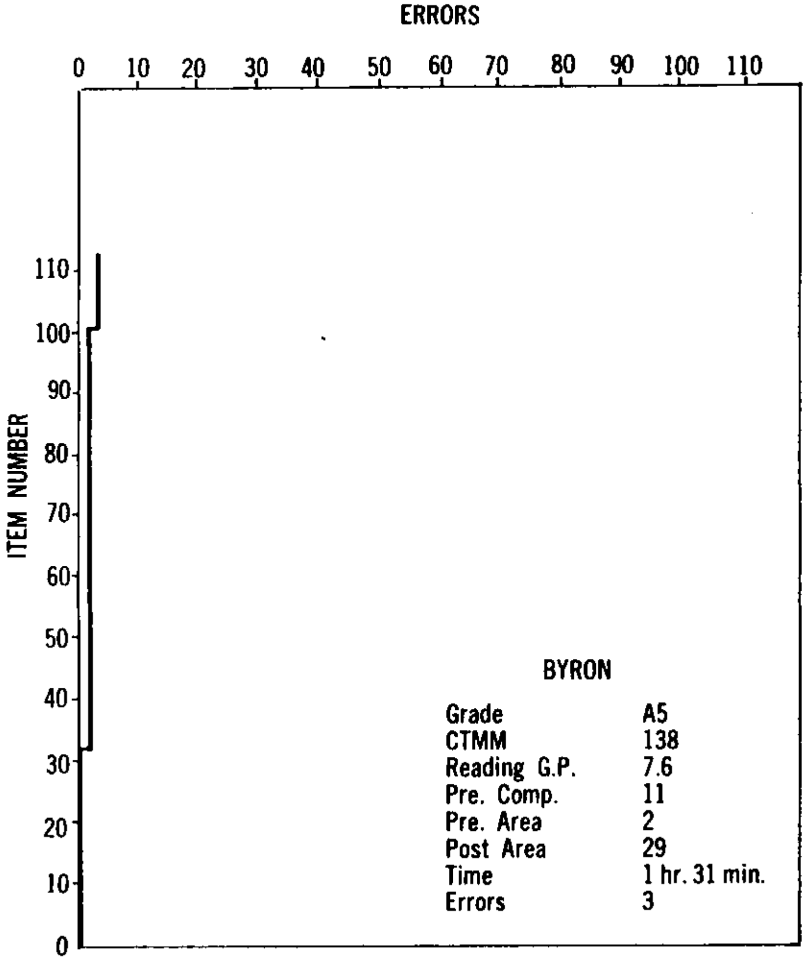


FIG. IV. Graphic record of the best performance

into two items, the first introducing the concept of square inch only. Item 25 could be rewritten as two or even three items.

The greatest misapplication of the principle of small steps occurred in the latter part of the program. For the group of 15 pupils who performed the item set, on the first 10 items there was a total of five errors on the first attempt (out of a possible 150). On succeeding sets of 10 items this number of first attempt failures

increased until it reached 66 for the last set of 10 items. The relatively high difficulty of the items in the latter part of the program appeared to result from the fact that these more complex items required a diversity of other understandings, abilities which were not tested for in the pretest. For example, although pupils were selected on the basis of their ability to divide, many failed to relate division to the process of successive subtraction. The program failed to provide adequate introduction or review of these concepts.

Faced with this type of problem, a programmer can either use a more adequate pretest to provide for better selection or he can add the additional items as required. When a great variety of item sets in arithmetic become available, the problem largely disappears. The prerequisites to any item set can be stated in terms of the successful completion of previous item sets. Conversely, the record obtained for a pupil with one item set might be used diagnostically to indicate what the next set, remedial, optional, or otherwise, might be.

3. A wider variety of items should be used for each new process. For instance, Items 25-28 should be supplemented with items which provide verbal statements as alternatives and more familiar illustrations of the problem. As another example, on many of the post-test questions asking for an area, subjects wrote only the correct number, failing to indicate the square units involved. Although six items had been presented in the program to reduce this kind of error, these items unfortunately were all stated in exactly the same way. The item stem called for an area, and among the alternatives one distractor listed only the correct number. Although completion items may be necessary to teach this type of behavior, better results in this program could probably have been obtained if, instead of the single form, a variety of multiple-choice forms had been used; e.g., use of "none of these" as an alternative; asking, "What is left out in this answer?"; or asking for the appropriate rule.

CONCLUSION

The use of multiple-choice items in automated teaching appears to have some effectiveness under these conditions in teaching understanding, as herein defined, even though the criterion was a free-answer test. Although the average pupil did not show as high a degree of competence on the post-test as expected, the program was also far more difficult than intended. Since there appeared to be a strong relationship between success on the program and gains on the post-test, before the limitations and advantages of the multiple-choice method used in this study may be assessed, the program should be revised to include smaller steps and a greater variety of items. Whether the best performers on the present program would have learned more from such a longer and simpler revision remains to be determined; to accommodate individual differences in ability, two or three versions of the program may be desirable.

SUMMARY

Fourteen elementary-school pupils responded individually to a set of 110 multiple-choice items in a teaching machine. The performance of each child was graphically recorded. Subjects performed significantly better on a test of understanding of areas of rectangles than did their matched controls who received no planned instruction on this topic. The principles of programming are discussed and illustrated. Suggestions given for the revision of the program, which appeared to be too difficult for most pupils, include the introduction of smaller steps and a greater variety of types of multiple-choice items.

An Evaluation of Textbooks In Terms of Learning Principles

ROBERT GLASER, LLOYD E. HOMME, and J. L. EVANS

A UBIQUITOUS aid to teaching is the textbook, and much effort has been invested in the improvement of the textbook with respect to such things as readability, legibility, and pictorial and graphic presentation. However, its basic method of transmitting information to the student has remained the same. Essentially, information is asserted in the form of declarative sentences on the assumption that learning will occur by some means. In the light of modern learning theory, the adequacy of this procedure can be seriously questioned. It appears reasonable that techniques can be developed which provide more specific control and adequate monitoring of the reader's behavior. This should result in the more effective transmission of knowledge. If textbooks can be constructed so that a student's learning is carefully guided, it is possible that present-style textbooks will become outmoded as teaching aids.

It is evident that even the best of standard textbooks fail to heed certain basic learning principles. Consider, for example, the educators' dictum that "active learning is superior to passive learning." In learning terminology, behavior must be evoked before it can be strengthened. Guthrie (1942) has made this point by saying: ". . . a student does not learn what was in a lecture or book. He learns only what a lecture or book causes him to do." Conventional textbook presentation does not insure that the appropriate behavior will be evoked or, if evoked, will be adequately strengthened. Consider together with the above notions the recent work of Skinner (1958c) who has concentrated upon the use of a machine to control the learning sequence. Skinner points out that a teaching machine has several impor-

---

This paper, originally presented at the meetings of the American Educational Research Association in February 1959, is one of a series of papers based on work on programmed learning sequences being conducted at the University of Pittsburgh. Dr. Homme and Dr. Evans are now with Teaching Machines, Inc., Albuquerque, New Mexico.

tant features: (a) The student must make overt responses; furthermore, he must compose his response rather than select it from a set of alternatives. (b) The student must pass through a carefully designed sequence of steps which are small enough to be taken correctly and which always move him closer to the behavior identified as the objective of instruction. (c) The machine must insure that these steps are taken in a prescribed order.

Although Skinner has concentrated upon the use of a "machine" to control the learning sequence, this does not preclude the possibility of other means of control. The teaching device which is being developed at the University of Pittsburgh can be designated as a "paper teaching machine" or as a "programmed textbook."

The programmed textbook is simply and easily described. Its external appearance will not differ from an ordinary textbook, but its interior is quite different. Each page consists usually of four or five panels. (See Table 1.) The sequence in which the panels are to be responded to is not from the top of the page to the bottom, as in a conventional textbook; but only one panel is "read" or responded to before the student turns the page. The student begins with the top panel on Page 1, i.e., Step 1; he responds to it, turns to Page 2 to get his answer confirmed on the top panel, then goes to the top panel of Page 3 for Step 2. He responds to it, confirms his answer by turning the page, and so on, to the end of the unit or chapter, where he is instructed to return to Page 1 and respond to the second panel on each page, and so on to the end of the chapter. Since each panel or step is numbered, no confusion results in this procedure.

With respect to the main functions of presently developed teaching machines as we have described them, the programmed text is clearly capable of fulfilling these functions which depend on the structure of the program rather than upon the technique of presenting it. This means that one is faced with the same problems in constructing a program, whether it be for a machine or a programmed textbook.

Following from both these lines of thought, i.e., textbook redesign and the development of teaching machines, the important basic problem obviously concerns the properties of the learning sequences that lead to the acquisition of knowledge. In the face of present relationships between the science of learn-

Table 1. AN EXAMPLE OF A PROGRAMMED TEXTBOOK; CONVERSION TO NUMBER BASES OTHER THAN 10—Page one

STEP 1: It has been said that the reason we use a **decimal** (from the Latin word for 10) system in our arithmetic is that primitive men probably counted on their fingers. The total number of fingers and thumbs on both hands is ..... . This number is called the **base** of our ordinary arithmetic.

STEP 3: There is nothing magical or sacred about the base 10 system we employ. In fact, conversion to bases other than 10 often provides important advantages. Some people have suggested that a base 12 system would facilitate computation, since 5 numbers **less** than 12 divide into it **evenly**. These 5 numbers are ..., 2 , ..., ..., ... In our base 10 system, on the other hand, only the numbers ..., 2 , and ... are less than 10 and divide it evenly.

STEP 5: Since we are familiar with the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, we will use them as much as possible while converting to other bases. For example, if we wanted to convert to the base 3, we would use the digits 0, 1, 2. If we wish to convert to the base 5, we would use the digits 0, 1, 2, 3, 4. If we wished to convert to the base 7, we would use the digits, ..., ..., ..., ..., ..., ..., ...

STEP 7: There is another base, the **binary** base, which is immensely important in the construction of modern high-speed computers. These computers use vacuum tubes and switches which can be either "ON" or "OFF." From this and from the word "binary," you might guess that the base of the binary system is .... .

Table 1 (Cont.). AN EXAMPLE OF A PROGRAMMED TEXTBOOK: CONVERSION TO NUMBER BASES OTHER THAN 10—Page two

STEP 1: Response Check

10

STEP 3: Response Check

1	1
3	5
4	
6	

STEP 5: Response Check

0
1
2
3
4
5
6

STEP 7: Response Check

2

Table 1 (Cont.). AN EXAMPLE OF A PROGRAMMED TEXTBOOK: CONVERSION TO NUMBER BASES OTHER THAN 10—Page three

STEP 2: If our primitive computers had had a different number of fingers on their hands, they may very well have had an arithmetic with a base other than 10. For instance, if they had had six fingers on each hand they might have developed an arithmetic with base ....

STEP 4: Since we use a base 10 system, we need 10 different numbers or symbols, which we will call digits, to write out our values. These digits are 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Notice the 0 (zero). You will need to list the 0 when writing out the digits of all bases. We always need the **same** number of digits as the base we are using, and one of these digits will always be "0." If we use a base 5 system we would need ...... digits.

STEP 6: We have seen that conversion to the base 12 might provide some advantages in computation, especially in division. However, if we use base 12, we will need ...... different digits. Since the digits 0-9 provides only 10 digits we must make up 2 new ones. (We wouldn't want to use 10 and 11, since these are made up of other smaller digits and are therefore not distinct.) You could make these up yourself, but for convenience, let us use $\cancel{9}$ for the digit after 9 and $\cancel{8}$ for the digit after $\cancel{9}$. Therefore, the digits for the base 12 are ......, ..., ..., ..., ..., ..., ..., ..., ..., ..., ..., ...

STEP 8: The binary system is base 2; therefore the digits we might use would be ......, ...

Table 1 (Cont.). AN EXAMPLE OF A PROGRAMMED TEXTBOOK: CONVERSION TO NUMBER BASES OTHER THAN 10—Page four

STEP 2: Response Check

12

Please return to Page One and go on to Step 3.

STEP 4: Response Check

5

Please return to Page One and go on to Step 5.

STEP 6: Response Check

12
0
1
2
3
4
5
6
7
8
9
~~9~~
~~9~~

Please return to Page One and go on to Step 7.

STEP 8: Response Check

0
1

ing and the art of teaching, the solution to this problem requires experimental investigation. Using the programmed textbook as instrumentation, such research appears quite feasible and fruitful. In such an endeavor three classes of variables are of interest: learning-sequence properties, knowledge (behavioral end-product) variables, and individual differences.

LEARNING-SEQUENCE PROPERTIES. This refers to the specific programming procedures or properties of the learning sequence which guide the student's behavior toward the acquisition of certain knowledges as he reads and responds to the material. This class of variables has been our main area of interest.

KNOWLEDGE (BEHAVIORAL END-PRODUCT) VARIABLES. Subject materials can be taught in order to result in different kinds of student behavior; for example, the learning of facts as compared with learning to solve problems or making practical applications. The properties and utility of different programming techniques and learning sequences for shaping up these different kinds of knowledge are important for study.

INDIVIDUAL DIFFERENCES. It is probable that success with certain learning sequences will be a function of variations in the intellectual levels and aptitude patterns of the learners. The differential effectiveness and characteristics of learning by means of programmed text procedures should be investigated for high-, low-, and exceptional-aptitude students to the degree that it would be feasible to construct differently programmed textbooks or different subsequences within a program to facilitate optimal learning.

Identification of the properties of such learning sequences (or the way in which learning is programmed) is certainly not a new endeavor for psychologists and educators. Many of the variables influencing the acquisition of knowledge through a programmed learning sequence come out of learning theory; others arise in the course of educators' progress in building effective learning sequences. Let us consider very quickly some of these variables here—those which have arisen in the course of our attempts at building programmed textbooks in various subject matters; namely, number theory, music fundamentals, statistics, and psychology. Let us now list an illustrative few.

1. RESPONSE UNIT SIZE (SIZE OF STEP). As presently conceived, a learning sequence consists of a series of units, each of which evokes a response from the reader. These successive advances can consist of larger or smaller approximations to the desired behavior. The optimal response unit size needs to be determined by objectively varying different-size response units in experimental programs. To accomplish this, different sequences can be constructed to teach the same concept with varying numbers of response evocations.

2. STIMULUS VARIATION (STIMULUS GENERALIZATION). A response unit provides a specific stimulus context to which the student responds. A response which is too specifically tied to a unique stimulus is said to be lacking in "adaptability" and "generalizability." In order to increase stimulus generalization, response panels must be varied in appropriate ways; the determination of the characteristics of this stimulus variation is an experimental question.

3. THE USE OF PROMPTS. In this context, prompts are defined as verbal stimuli which increase the probability that a given verbal response will occur. A basic notion in learning sequences is that sufficient prompts are always present to evoke a correct response. As the learner's repertoire develops from an elementary to an advanced stage, the nature and frequency of the available prompts change. Certain features of prompting are of great interest for study.

4. RESPONSES EVOKED (RULES FOR OMISSION). Few if any rules are available for the kinds of responses that should be evoked by the omissions in a response panel. Presumably the objective of a learning sequence is to shape up certain verbal (including symbolic) behavior. If this desired behavior can be satisfactorily specified, then the question is, What kinds of verbal responses need to be evoked in order to produce it most effectively? Behaviors which are acquired primarily as a basis for transfer probably require different evocation histories.

5. FORM OF A RESPONSE. Practically all of the learning programs prepared to date have employed written overt responses to a response panel. The form or modality of the response may need not be so restricted. If the responses for a particular subject matter can be accomplished without overt activity, i.e., implicitly, this increases the speed through a learning sequence. The primary consideration here is the response modality required to produce the desired behavioral outcomes. If mathematical manipulation is the desired behavior, overt responding is probably necessary. In any case, the results produced by utilizing different kinds of responses in a learning sequence, e.g., spoken, written, or implicit responses, are a matter for study. Particular kinds of responses may be especially necessary for individuals with handicaps in various modalities.

6. RESPONSE "STYLE". Just how should the student respond? Skinner emphasizes the great importance of "composing" a response and that prompts and stimulus supports be presented so that the appropriate response will be evoked. The response of the student is then confirmed by the correct response appearing in the response panel. Responses may also be evoked in other ways; i.e., by selecting from response alternatives in multiple-choice fashion or by a combination of composing and multiple choice. The important question seems to be the influence on the desired end-product knowledge. Does multiple-choice responding result in poor essay writing or in inadequate problem solving or creative proficiency? Does composing a response permit more flexibility in shaping it? These questions need to be answered in the experimental investigation of learning sequences.

Other learning-sequence variables that we have considered are: word-to-response ratio, response "position in a frame," incidental learning, discrimination-training procedures, spacing of practice, and error threshold. [Results of preliminary investigation of some of the properties of learning sequences, summarized in the original version of this paper, are reported in the following paper of this volume.]

A Preliminary Investigation of Variation In the Properties of Verbal Learning Sequences of the "Teaching Machine" Type

J. L. EVANS, ROBERT GLASER, and LLOYD E. HOMME

As SKINNER (1958c) has pointed out, the success of teaching machines is dependent upon the material or "program" used in them. Paradoxically, however, the technology for the construction of machines seems to be growing more rapidly than the technology for the construction of the programs. Perhaps this is so because it is easier to write specifications for the machine and because it is more intriguing to build machines than it is to write programs.

Such programs—or to use a more descriptive phrase, verbal learning sequences—are of critical importance in the success of ventures in automated learning. Work with such sequences has led us to two conclusions:

1. Verbal learning sequences and their properties provide an interesting field for psychological investigation in their own right.
2. The properties of such sequences can be studied experimentally without actually employing a "hardware" machine.

The last statement is not meant to imply that the role of the machine is unimportant. The determination of just what functions the machine should provide beyond such obvious ones as automatic recording and prevention of "cheating" by the student has yet to be made. Whether or not the machines have intrinsic properties which cannot be matched by non-machine techniques awaits controlled experimental comparisons of machine and non-machine presentation of sequences. In any event, it is the verbal sequences themselves with which we wish to deal in this paper.

Studies of verbal learning sequences can be divided conveniently into two classes. In the first case, we may be interested in varying some properties of a given learning sequence and assessing the effects of such variations on performance. In the second case, we may wish to evaluate the relative effectiveness of sequence learning compared with more conventional processes such as textbook presentation. This second group of "comparative" studies may be somewhat premature in light of our meager knowledge about the variables relevant to the production of effective learning sequences, but it is very tempting to try to evaluate a learning sequence as soon as it has been written. In this respect, it is perhaps surprising that learning sequences, constructed without the use of many specific rules or experimental evaluation, have shown up as well as they have. Studies of both kinds described above have been conducted at the University of Pittsburgh.

The learning sequences which we used in our work consisted of a series of short written steps, each of which called for one or more written responses on the part of the subjects. We presented our steps either on separate index cards or in a "programmed textbook." The programmed textbook is simply and easily described. . . . [See the preceding paper by Glaser, Homme, and Evans.]

Moving on to the first class of experiments, those in which properties of the sequences themselves are varied, two preliminary studies have been undertaken. In Experiment 1, summarized in Table 1, a sequence was utilized which was designed to teach "conversion to number bases other than 10." As originally written, this sequence had 51 steps. Sequence writing typically involves much repetitive and transitional material which is largely omitted by textbooks. The question was raised whether or not it would be possible to produce the same level of performance with smaller amounts of such redundant and transitional material. The effect of editing out such material should be to get the learner over the same distance, but in larger "steps," since fewer separate steps were being taken and fewer written responses were required. Steps were actually removed from the program, with care being taken to leave in the essential

---

This paper was first presented at the April 1959 meetings of the Eastern Psychological Association. It is based on work on programmed learning being conducted at the University of Pittsburgh. Dr. Evans and Dr. Homme are now with Teaching Machines, Inc., Albuquerque, N.M.

Table 1. SUMMARY OF EXPERIMENTS

EXP. NO.	TOPIC	GROUPS	N	TREATMENTS	RESULTS
1.	Conversions to number bases other than 10	A	5	Used 30-step sequence	1. Small-step sequences produced significantly better performance ($p < .05$) on both immediate and retention tests than did the shorter large-step sequences. 2. Small-step sequences produced significantly fewer response errors ($p < .05$) during the learning sequence despite increased opportunities for error. 3. Small-step sequences in general took more time than large-step sequences. 4. Group D performance was slightly inferior to Group C.
		B	5	Used 40-step sequence	
		C	5	Used 51-step sequence	
		D*	5	Used 68-step sequence	
2.	Fundamentals of music	W	5	Instructed to write down answers to each step	1. Group N took less time to complete sequence (NS).* 2. Group N displayed greater mean performance gains (NS). 3. Group W displayed greater variability in performance scores (NS).
		N	5	Instructed not to write down answers to each step	
3.	Measures of central tendency and variability	S	16	Material presented in learning-sequence form	1. Group S had a higher mean performance test score (NS).* 2. Group T displayed significantly more variability in performance measures ($p < .05$).
		T	17	Material presented in text form	
4.	Measures of central tendency and variability	S	5	Material presented in learning-sequence form	1. Group S had higher mean score on performance test (NS). 2. Group T displayed significantly more variability on performance measure ($p < .01$).
		T	8	Material presented in text form.	
5.	Fundamentals of music	S	5	Material presented in learning-sequence form	1. Group S had higher mean performance scores ($p < .05$).
		A	5	Material presented in text form	

*Group D was added three months after initial experiments on Groups A, B, and C.

*Not Significant

points of the development of the subject matter. Also, items were *added* to the original sequence so that, in all, *four* sequences with 30, 40, 51, and 68 steps respectively were constructed. Four independent groups of five *Ss* each, who were graduate education students taking psychology courses, were given the sequences. Upon completion of a sequence performance, measures were taken by means of written test. Results indicate that "smaller" steps, i.e., those using more steps in a sequence to cover the same subject matter, were associated with significantly fewer errors on immediate and delayed performance tests. Smaller steps also were associated with less time per step and fewer errors in the course of learning. However, as suggested by the results obtained with Group D, in Table 1, it appears that there is a limit beyond which it does not pay to extend the number of steps. On the test immediately following completion of the sequence, the performance of Group D was similar but slightly inferior to Group C. In general, increasing the number of steps resulted in more time being taken to complete a sequence.

The second study involving properties of the learning sequence used a sequence on "fundamentals of music." In this experiment two groups of undergraduate psychology students were used. The first, or "W," group went through the sequence in the usual manner, producing one or more written responses at each step. The second, or "N," group proceeded through the sequences, but were instructed *not* to make an overt written response. The analyses summarized in Table 1 indicate that members of the Group N (who did not write their answers) did better, but not significantly so, on the performance test. Group N also took less time to complete the sequence. This finding seems to demand some re-examination of procedures which require overt written responses on the part of subjects, as is typically the case in teaching machine work.

The next three studies are concerned with making an experimental evaluation of sequence learning and textbook presentation of the same material. In Study 3, one group of 17 undergraduates in psychology was given the material covering 10 pages of a standard statistical text (Edwards, 1946). A second group of 16 was given a learning sequence, or programmed text, over the same material. An independently constructed multiple-choice test was then administered. Group S, which received the learning sequence, had a higher mean performance score, but this difference was not significant. It is interesting to note, however, that the learning-sequence group displayed significantly less variability in their scores. Examination of the results of this study indicated that the multiple-choice test used was too easy to discriminate performance properly. A revised test was therefore constructed which included a "short answer" test over the material. In Experiment 4, two new groups were given the same experimental material. Again, the learning-sequence group had higher performance scores, which approached statistical significance more closely this time but did not reach it. Also, the learning-sequence group displayed significantly less variability in performance measures, as before.

Experiment 5 was conducted in the same manner as Experiments 3 and 4, but here the subject matter was "fundamentals of music." In this study the learning-sequence group did perform significantly better ($p < .05$) on a test of identification of musical notation. The learning-sequence group also displayed less variability in its scores. Since the results of the three independent studies comparing learning sequence and text presentation of materials all gave some indication of tending toward statistical significance, the three probability values were combined following a procedure outlined in Edwards (1954, p. 392-93). The resulting Chi-square value of 13.05 for 6 degrees of freedom was significant at the .05 level. It should also be pointed out that the learning-sequence groups typically take significantly more time to complete their sequence than do the textbook groups.

In Summary:

1. Learning sequences consisting of smaller steps are associated with significantly better immediate test performance, better retention, and fewer response errors in the course of learning.
2. No significant differences were found between *Ss* who were instructed to make overt written responses to each step in the learning sequence and those *Ss* who were instructed *not* to make written responses.
3. In general, *Ss* receiving learning sequence made higher achievement score and exhibited less variability in performance than did *Ss* receiving conventional textbook presentation of the same material.

In conclusion, it appears that the properties of verbal learning sequences or "programs" can be investigated experimentally without employing a "hardware" teaching machine, and it is suggested that whatever techniques and principles should emerge from such studies may find immediate practical application in the field of verbal behavior.

Results of An Initial Experiment In Automated Teaching

J. E. COULSON and H. F. SILBERMAN

INTEREST in automated teaching has increased greatly in the past 10 years, as evidenced by the number of papers on the subject given in recent psychological and educational conferences and by the variety of teaching machines now available for research or commercial purposes. This new interest has been stimulated by the critical shortage of skilled teachers throughout the country and by the belief of many psychologists and educators that automated teaching devices can be a valuable tool for teachers in both academic and industrial situations. For some learning situations the teaching machine may surpass its human counterpart, particularly where precise control of the timing and contingencies for student reinforcement is critical.

A teaching machine may be defined generally as a device which has the following operational characteristics:

1. It presents problem materials (questions) to the student.
2. The student is required to respond to the materials by overt behavior.
3. The machine provides the student with knowledge of results of his behavior, usually immediately following each response. This "knowledge of results" consists of telling the student whether his response was correct or of providing the student with information from which he can evaluate his own response.

Within the generic definition given above there are many dimensions along which one teaching machine may differ from another. The purpose of the present study is to investigate three of these dimensions:

1. *Student response mode.* The relative advantages of the multiple-choice and the constructed-response modes of operation have been debated at length in the literature on teaching machines. S. L. Pressey (1950) has been a leading proponent of the multiple-choice mode, while B. F. Skinner (1958c) has gathered many adherents to the position that constructed-response items are superior. Response mode was included in the present study both because of its interest as a variable per se and because of the possibility that it would interact significantly with the other two variables.
2. *Size-of-item step.* Most investigators in the field of automated teaching agree that optimal learning occurs when items presented to the student give the student a large degree of support (Homme and Glaser, 1959b; Porter, 1958c; Skinner, 1958c). According to these investigators, the items should be worded and sequenced in such a way that the student finds it simple to proceed from one item to the next and thus receives a large percentage of positive reinforcements. In the present study an attempt was made to evaluate experimentally the extent of the purported superiority of small-step conditions. Certain items were eliminated from one item set to create another set containing fewer items. Eliminated items covered essentially the same subject material as other items in the set and introduced no new concepts. The derived set may be operationally defined by this process as having larger step size than the original set.
3. *Item sequence control.* The great majority of existing teaching machines present subject-material items in an essentially predetermined sequence. Relatively little speculation and even less actual experimentation have been devoted to the possible advantages of a sequence determined by the behavior of the student.

This is an edited version of a report prepared at the System Development Corporation, Santa Monica, California. Extensive statistical tables and appendices including experimental materials and instructions to the subjects in the original report have been deleted or abridged by the editors of the present volume. The authors wish to acknowledge the assistance of the following persons: C. M. Bertone, S. Cooney, R. E. Dear, W. R. Ellis, Jr., K. Esgate (Santa Monica City College), C. Fuller, T. E. Henderson, B. T. Jensen, S. L. O'Connell, B. Patton, C. D. Parkinson, M. S. Rogers, Matilda Sims, L. Staugas, W. Thomas (Santa Monica City College), and R. J. Tuber. This study is also reported in the *Journal of Educational Psychology*, 1960.

It is true that most teaching machines in use today have some limited flexibility, in the sense that a student must answer a question correctly or be shown the correct answer before the next question is presented. Being shown the correct answer or selecting the correct answer after several tries does not insure, however, that the student has actually acquired the necessary knowledge. If he really does understand the concept involved, it is probably inefficient to present many additional questions covering the same concept. If the student does *not* understand the concept, on the other hand, he should not be presented with more complex questions based upon the original concept, since he is not properly prepared to answer such questions.

A truly adaptable teaching machine should be capable of discriminating critical aspects of the student's response behavior and selecting subsequent instructional items on the basis of this behavior. Other variables which a teaching machine might modify in response to student behavior include the amount and type of knowledge of results presented to the student and the number of "cues" given to help the student answer the problem materials. Such a teaching machine would act much more like a human tutor than do most machines available today.

In the present study a simple type of sequence flexibility, which will be called "forward branching" or merely "branching," was compared with the conventional predetermined sequence. The branching procedure (described in more detail in the following section, "Method") was intended to allow a student who has already learned a concept to skip over certain other items covering this same concept. It was hypothesized that students trained with the branching procedure would learn as well as students trained without branching and would learn in a shorter time due to the omission of some of the items.

METHOD

Apparatus and Procedures

The experiment consisted of three phases: (a) a training session for each of 80 subjects with a teaching machine operating under one of the eight different teaching procedures being compared, (b) a criterion quiz given to each subject immediately following his training session to determine how much the subject learned in the training session, and (c) the same criterion quiz given approximately two weeks after the training sessions to all students in the school classes from which the experimental subjects were drawn. This last phase was intended to provide a measure of retention in the subjects as well as a comparison of subjects with other students who did not have the experimental training session.

Although the training session for any single subject required less than two hours, the sessions for all subjects extended over a period of one week. During this week the training sessions replaced regular classroom studies for the experimental subjects, while the control group continued to receive its normal instruction. Subject material for the classroom instruction during this week, and up to the final administration of the criterion test, was not closely related to the material taught with the teaching machines.

Six identical sets of equipment were used in the experimental training sessions. The sets actually constituted *simulated* teaching machines, since for purposes of convenience, human experimenters were used in place of automatic control mechanisms [cf. Rothkopf, p. 327.—Eds.]. Each set consisted of a wooden screen, a number of push buttons and lights used for communication between experimenter and subject, a deck of 5" x 8" cards containing instructional items, and four sheets ("panels") with information relating to the instructional items. The wooden screen had a window in its center in which the experimenter placed cards so the instructional items could be seen by the subject. A holder on the experimenter's side of the window supported the cards and also prevented the subject from observing the experimenter through the window at any time. The information sheets were attached to the subject's side of the screen for all experimental training sessions and could be used by the subject as an aid in answering the instructional items; subjects could not use the sheets while taking the criterion quiz, however.

Each of the six sets of experimental equipment could be operated under any one of eight teaching procedures. These eight teaching procedures represent the eight combinations of three experimental variables, each variable having two possible values, as follows: (a) response mode (multiple choice versus constructed response), (b) size of steps between successive items to be taught (small steps versus large steps), and (c) type of item sequence control (branching versus nonbranching). Thus each subject was run under one of the following conditions:

1. Multiple choice, small steps, no branching
2. Constructed response, small steps, no branching
3. Multiple choice, large steps, no branching
4. Constructed response, large steps, no branching
5. Multiple choice, small steps, branching
6. Constructed response, small steps, branching
7. Multiple choice, large steps, branching
8. Constructed response, large steps, branching.

Subjects and experimenters were randomly assigned to the eight experimental conditions, within the restriction that 10 subjects were trained under each condition.

Printed instructions were read aloud to each subject before the start of the training session. These instructions described the task to be performed by the subject and informed him that both accuracy and speed of learning were important. [The instructions were reproduced verbatim in appendices in the original report.—Eds.]

CONSTRUCTED-RESPONSE VERSUS MULTIPLE-CHOICE OPERATION. In the constructed-response mode of operation there are two push buttons on the subject's side of the screen and two small lights on the experimenter's side. The experimenter places in the window of the screen a card containing a constructed-response item. This item consists of a statement with one or more blank spaces representing words to be filled in by the subject. The subject writes his answer on a slip of paper and then taps lightly on the screen to indicate that he has made his response. Upon hearing the tap, the experimenter removes from the window a cardboard mask, revealing to the subject the correct answer to the item. The subject compares his own answer against the correct answer and uses one of his two push buttons, marked "right" and "wrong" respectively, to activate the corresponding light on the experimenter's side. This indicates to the experimenter whether or not the subject got the item correct.[1] The experimenter removes this card and goes on to the next card, continuing until he has finished the deck. He then repeats the entire process but this time uses only the items missed by the subject on the first trial. The procedure is continued through successive trials until the subject has answered every item correctly.

In the multiple-choice mode there are five push buttons on the subject's side of the screen, activating five lights on the experimenter's side. These buttons are used by the subject to indicate his answer to the multiple-choice items presented to him. In addition, there are two push buttons used by the experimenter to activate lights on the subject's side, telling the subject whether he has selected a correct or incorrect response. Multiple-choice items consist of statements with blank spaces representing missing words. Below the statement are listed two to five possible choices, one of which is to be selected by the subject to answer the item. This selection is accomplished by the subject's pressing the push button having the same letter label as the answer which he wishes to indicate. The various choices are presented to the subject simultaneously with the item, there being no mask of the sort used with constructed-response items. The subject must continue to respond to an item until he has selected the correct answer, upon which the experimenter replaces that item with the next in the deck. After he has completed the first trial (e.g., the first run through the deck), the experimenter continues with additional trials as required, each time using only the cards that the subject *missed* in his first response on the preceding trial.

"SMALL STEPS" VERSUS "LARGE STEPS." Some subjects were trained with instructional decks containing many (104) items. These decks are called "small-step" decks, because they supposedly require little effort on the part of the subject to answer a particular item once he has gone through the preceding items. Other subjects were trained with "large-step" decks containing fewer (56) items and ostensibly requiring more effort to answer successive items. From the standpoint of the operation of the teaching machines, there are no procedural differences between small-step decks and large-step decks, except in the number of items presented to the subject.

---

[1] An informal check was made to determine the dependability of the subjects' evaluation of their own responses. Several randomly selected subjects were observed during training; and after training, their slips were compared with the correct answers. Results of the check indicated that subjects were generally accurate in their evaluations.

"BRANCHING" VERSUS "NONBRANCHING." This variable determines the sequence in which items are presented to the subject. In the "nonbranching" mode, the experimenter simply goes through all the cards of the deck in order. In the "branching" mode, certain items are removed from the deck if the subject answers certain other items correctly on the first try. The items to be removed are taken out of the deck by a second experimenter, so they are never presented to the subject. Thus in the branching mode the exact number and sequence of items are not fixed but depend on the performance of the subject. A more complete description of the criteria for omitting items is given in the following section, "Subject Material."

PERFORMANCE MEASURES. The following performance measures were recorded for each subject, regardless of mode of instruction:

1. Time required to finish training
2. Score on multiple-choice portion of criterion test
3. Score on constructed-response portion of criterion test.

Subject Material

The eight item sets (decks) and the three instructional sheets used in the experiment were based on a portion of a college course in elementary psychology used at Harvard University. The first item set obtained was the constructed response, small-step decks, without branching. This set consisted of the first 104 items from a larger series of items developed at the Harvard University Psychological Laboratory [see Holland, 1960]. The remaining seven sets represent combinations of the following transformations of the original items:

1. A number of possible alternative answers were provided for each item to make the multiple-choice decks.
2. Certain items were eliminated from the original set of 104 items to make the large-step decks of 56 items. The items eliminated were judged to be largely redundant, in the sense that they related to concepts already covered by other items and added little extra information.
3. To make the branching decks, coded instructions for the branching procedure were marked on the backs of certain cards. These instructions indicated which cards were to be eliminated from the deck during the experiment if the subject correctly answered the marked cards. For convenience, a marked card will be referred to as a "branch card"; the cards to be eliminated when a branch card is correctly answered will be called "conditional-skip" cards. In general, a branch card was the first of several items in the deck covering a particular concept, while the associated conditional-skip cards consisted of the remaining items covering essentially the same concept. Small-step branching decks included 12 branch cards and 39 conditional-skip cards, while large-step branching decks included 12 branch cards and 22 conditional-skip cards. The three instructional sheets, or "panels," used in the study covered most of the major concepts presented in the instructional items but in the form of definitions and brief descriptive materials.

Criterion Test

The criterion test consisted of 36 questions, of which 19 were constructed response (fill-in) and 17 were multiple choice. All of the questions were based upon the material contained in the original (constructed response, small step, without branching) set of instructional items. The person who prepared the criterion items had not seen the other seven item sets prepared for instruction; conversely, the persons preparing the alternate item sets had not seen the criterion questions.

None of the criterion questions were duplicates of instructional items, though many of the same words were used as correct responses. Most of the criterion questions were of the application type, in which a situation is presented and the respondent identifies the principle involved or attempts to explain or predict an outcome.

Reliability of the criterion test, as estimated by the application of Kuder-Richardson Formula 20 to the scores of the experimental subjects, was as follows: whole test, .89; constructed-response portion, .85; multiple-choice portion, .79. Test-retest correlations for experimental subjects were: .81 for the whole test, .79 for the constructed-response portion, and .58 for the multiple-choice portion.

Experimental Subjects

The 80 experimental subjects and the 104 members of the control group were taken from beginning psychology classes in Santa Monica City College. They had been in the psychology course for about one month . . . covering some general background in the area of psychology, but had not been given specific instruction in any material closely related to the subject material used in the experiment.

RESULTS

1. A psychology pretest was given to all experimental subjects and also to the students who had not been selected for the experiment. No significant difference was obtained between the 80 experimental *Ss* and the 104 nonexperimental *Ss* using a "t" test.

Table 1. MEANS AND STANDARD DEVIATIONS FOR EXPERIMENTAL AND CONTROL GROUPS ON PRETEST AND CRITERION SUBTEST

SOURCE OF SCORES		EXPERIMENTAL GROUP		CONTROL GROUP	
		Mean	SD	Mean	SD
PRETEST		45.5	8.0	44.4	7.7
FIRST ADMINISTRATION OF CRITERION TEST	Multiple choice	11.6	2.5	10.5	2.6
	Constructed response	13.5	3.9	8.6	3.6
	Total criterion	25.1	5.9	19.1	5.2
SECOND ADMINISTRATION OF CRITERION TEST	Multiple choice	12.4	2.5		
	Constructed response	13.4	3.4		
	Total criterion	25.8	5.3		

2. After the experimental training period a criterion test, covering applications of the concepts that were being taught, was administered to experimental and nonexperimental groups. The criterion test consisted of two subtests, one with multiple-choice items, the other with constructed-response (fill-in) items. Table 1 gives the mean scores and SD's for the experimental and control groups. Maximum possible scores on the criterion were 19 and 17 for the constructed-response and multiple-choice subtests respectively. The experimental group was superior to the control group on the total criterion and on both the multiple-choice and constructed-response subtests. Differences were significant at the .01 level.

3. The criterion test was readministered to the experimental *Ss* three weeks after the first administration as a measure of retention. Means and SD's for the experimental group retest are given in Table 1. No significant difference was obtained on the total criterion (multiple choice plus constructed response), when the mean of the first administration was compared with the mean of the second administration. On the multiple-choice criterion subtest, however, retest scores were significantly higher (.01 level) than original test scores. This improvement in performance may be attributable to the fact that students became acquainted with the criterion questions, although it is difficult to explain why a similar improvement did not occur on the constructed-response subtest.

The evidence from these first three analyses indicates that the automated teaching materials used with the experimental subjects resulted in significant learning of the concepts taught and that this learning is retained at least for a three-week period. These findings do not mean that the experimental *Ss* exceeded groups receiving conventional classroom instruction on the same materials, since the control *Ss* were not being taught the same concepts as the experimental group. The purpose of the experimental-versus-control-group comparision was to determine whether *any* learning was taking place in the experimental group over that which might have occurred without such training.

4. The correlations between pretest scores and criterion scores were as follows: correlation between pretest and constructed-

response scores = .411 (.01 level of significance); correlation between pretest and multiple-choice scores = .404 (.01 level); correlation between pretest and total criterion = .442 (.01 level).

5. Results of covariance analysis of the constructed-response criterion measures for the eight experimental groups are shown in Table 2. [Table IV in the original report.] Pretest scores were used as a control variable in this analysis because of the significance of the correlation between pretest and constructed-response criterion scores. A significant main effect was found for the size-of-step factor (.05 level of significance), in the direction of higher scores for small-step trainees. A significant interaction was also found between the mode of response and branching variables (.05 level). The assumption of homogeneity of regression was tested and found to be satisfied.

Table 2. COVARIANCE ANALYSIS OF CONSTRUCTED-RESPONSE CRITERION SCORES USING A PSYCHOLOGY PRETEST AS A CONTROL

SOURCE OF VARIATION	ADJUSTED SUM OF SQUARES	df	ms	F
a. Mode of response	29.22	1	29.22	2.48
b. Size of step	46.99	1	46.99	3.99*
c. Branching	1.27	1	1.27	.11
AB	11.05	1	11.05	.94
AC	53.28	1	53.28	4.52*
BC	18.34	1	18.34	1.56
ABC	2.23	1	2.23	.19
Within cells	846.88	71	11.93	
Total	1009.26	78		

* .05 level

Table 3 [Table V in the original report] gives observed mean constructed-response scores for comparison of the effects of the three experimental variables (response mode, step size, and branching procedure). Mean constructed-response criterion scores are also given for the four combinations of response mode and branching procedure, since this interaction was found significant in the covariance analysis. The right-hand column of Table 3 shows the constructed-response criterion means adjusted for pretest scores.

Table 3. OBSERVED AND ADJUSTED MEAN SCORES ON CONSTRUCTED-RESPONSE CRITERION SUBTEST

GROUPS (TRAINING MODE)	OBSERVED MEANS ON CONSTRUCTED-RESPONSE CRITERION	CONSTRUCTED-RESPONSE MEANS ADJUSTED FOR PRETEST SCORES
Constructed response	14.05	14.14
Multiple choice	12.95	12.86
Small step	14.20	14.31
Large step	12.80	12.69
Branching	13.65	13.21
No branching	13.35	13.79
Branching with constructed response	13.55	12.93
Branching with multiple choice	13.75	13.50
No branching with constructed response	14.55	15.35
No branching with multiple choice	12.15	12.92

6. Covariance analysis of the multiple-choice criterion measures of the subjects in the eight experimental groups was also performed with psychology pretest scores as a control variable. The results of this analysis yielded no significant main effects or interactions among the experimental groups.

7. Covariance analysis of the total criterion measures (multiple choice plus constructed response), using the psychology pretest as a control variable . . . [showed] no significant main effects or interactions . . . among the experimental groups.

8. An analysis of variance of differences among the eight experimental groups with respect to the amount of time taken to complete training was also performed. All three main effects are significant at the .01 level. The constructed-response training condition took more time than the multiple-choice condition; the small-step condition took more time than the large-step condition; and the nonbranching condition took more time than the branching condition. Interactions were not significant. The mean scores on the time criterion are shown in Table 4 [Table X in the original report].

Table 4. MEAN SCORES ON TRAINING-TIME CRITERION

GROUPS	MINUTES
Constructed response	54.4
Multiple choice	44.4
Small step	57.3
Large step	41.5
Branching	43.8
No branching	55.0

DISCUSSION

Constructed-Response Criterion Data

SIZE-OF-STEP VARIABLE. The finding that treatment groups given more items (small steps) learned more than groups receiving few items is in accord with B. F. Skinner's emphasis on the importance of small steps in writing instructional item sets. The superiority of small steps in total amount learned must be weighed, however, against the significantly greater training time required by small-step trainees than by large-step trainees. The best choice of item step size for any applied teaching situation will be determined in part by the extent to which required training time is a critical consideration.

BRANCHING VARIABLE. In the comparison of large item steps and small steps it was found that small-step trainees, who were given more items, learned more than large-step trainees. In view of this finding, it is notable that branching, with its consequent reduction in total number of items presented, did not differ significantly from nonbranching in amount taught. One possible explanation lies in the fact that the difference in the number of items given to branching and nonbranching trainees was less than the difference between large- and small-step trainees. A more important factor, however, may have been the way in which items were eliminated in the branching procedure. An item was skipped only after the trainee had demonstrated some knowledge of the concept taught by that item, with the result that every major concept was covered by at least one item. At the same time, the branching procedure made possible a significant decrease in required training time, as compared with the nonbranching procedure.

When both the amount learned and the required training time are considered, the branching procedure appears to offer an over-all advantage over nonbranching. The present investigators feel, moreover, that branching or other types of control flexibility offer even greater potential for future improvement in teaching machine effectiveness. In the study reported here, limitations due to human simulation of the teaching machine control mechanism placed heavy restrictions on the branching procedure. The greatly simplified skipping procedure resulted in a significant advantage in training time but not in amount learned. More complex forms of control flexibility, using some type of automatic equipment, may prove superior to predetermined sequence control in both amount learned and speed of learning.

RESPONSE-MODE VARIABLE. Under the conditions of the experiment reported here, student response mode did not significantly affect the amount learned by the students. Since required training time was significantly less for multiple-choice trainees than for constructed-response trainees, the over-all advantage appears to be with the multiple-choice mode. It is possible, of course, that the results would be different under other experimental conditions, and further research in the area may prove valuable.

An important consideration in the evaluation of response mode for applied purposes is the ease with which a particular mode can be incorporated into an actual teaching machine. This practical consideration must be weighed along with the teaching effectiveness of the response mode. The use of the multiple-choice mode, rather than the constructed-response mode, greatly simplifies the design of any teaching device which is intended to evaluate student responses automatically.

RESPONSE-MODE–BRANCHING INTERACTION. The significant interaction between response mode and branching procedure can probably be attributed to the combined effects of two factors, neither of which is sufficiently powerful to have statistically significant effect when taken separately. The first factor contributing to the interaction is the number of items presented to the subjects. By this hypothesis, the constructed-response trainees learned more without branching than with branching because the branching procedure caused a reduction in the total number of items. The second factor contributing to the interaction effect was the relation between the response mode required in the training session and that required in the criterion test. As previously noted, all the criterion scores being compared here were obtained from the constructed-response portion of the criterion test. This might well be expected to give an advantage to the constructed-response trainees due to generalization decrement on the part of the multiple-choice trainees.

Multiple-Choice Criterion Data

Although the mean of the experimental *Ss* on the multiple-choice criterion significantly exceeded the control-group mean, no significant differences were obtained among the different experimental groups. Differential effects of the various experimental treatment combinations on the multiple-choice criterion appear to have been masked by a larger effect common to all experimental groups. This masking error variance can probably be attributed to the nature of the multiple-choice criterion subtest. Of the 17 multiple-choice items, nine had four alternatives, two had three alternatives, and six had only two alternatives. The probability of obtaining correct answers on a chance basis could account on the average for approximately one-third of the items. The largest portion of the variance of the scores on the multiple-choice test may be due to this chance factor. Students of the sort used in the study are usually highly skilled in picking out subtle specific determiners by which they can eliminate implausible alternatives on grounds other than factual content. These skills, which together with day-to-day fluctuations in individual performance are treated as systematic variance in the Kuder-Richardson method of estimating reliability, contribute to overestimation of the effective reliability of the test. In this case, the Kuder-Richardson reliability estimate was .79. A more realistic estimate of reliability is given by the multiple-choice criterion test-retest correlation, which was only .58.

An alternative to the "masking" interpretation is that there was actually no difference in effectiveness among the experimental treatments. The results on the constructed-response criterion and the operational differences among treatments, however, tend to place doubt on this second hypothesis. Although not significant, the direction of differences among the means of the experimental groups tend to support the results obtained on the constructed-response criterion.

Total Criterion Test Data

The results of the analysis of the total criterion follow from the results obtained on the analyses of the two subtests. Error variance in the multiple-choice portion of the total criterion appears to have masked intergroup differences so that these differences are not significant. The covariance F test approaches but does not reach significance at the .05 level for both the size-of-step factor and the mode of response—branching interaction. The direction of the means also agrees with differences obtained on the analysis of scores on the constructed-response criterion.

Time Criterion

Differences in required training time are generally in the expected direction. Small-step groups and nonbranching groups took longer than large-step and branching groups because they were required to answer more questions. The fact that constructed-response groups took longer than multiple-choice groups indicates that the composition and writing of answers are more time-consuming than the recognition of a "correct" solution among several alternatives, even though multiple-choice trainees were sometimes required to make several selections before choosing the right answer.

SUMMARY

Eight groups of 10 junior-college students were given an experimental training session with manually simulated teaching machines, each group being taught with a different mode of teaching-machine operation. The three independent variables were student response mode, size of steps between successive items, and sequencing (branching) procedure. A written criterion test was given to all subjects immediately after the training session and again three weeks later. The same criterion test was given to a control group which had no training with the teaching machine but which came from the same school classes as the experimental subjects. Dependent variables were the required teaching-machine training time and scores on the criterion test. Scores on a pretest were used as a control variable in a covariance design. Major results were as follows:

1. Use of the simulated teaching machines led to significant learning by the subjects, as determined by comparison with the control group.
2. The multiple-choice response mode took significantly less time than the constructed-response mode. No significant difference was obtained between response modes on the criterion test.
3. Small item steps required significantly more training time but also yielded significantly higher test scores than large item steps on the constructed-response criterion subtest.
4. The branching conditions required less training time than nonbranching but were not significantly different on the criterion test. A significant interaction was obtained between the mode of response and branching variables on the constructed-response criterion. This interaction resulted from a high mean criterion score obtained by the constructed-response–nonbranching group.
5. No significant differences were obtained among the experimental groups on the multiple-choice criterion subtest or on the total (multiple-choice plus constructed-response) criterion test.

A Study of Teaching Machine Response Modes

EDWARD B. FRY

THIS STUDY attempted to cast some light upon an important dichotomy in the field of teaching machines. The basic question investigated was: "In a teaching-machine situation, which mode of responding, multiple choice or constructed (write-in short answer), promotes more effective learning?"

A teaching-machine situation is defined as having the following characteristics, which are also found in most automatic teaching devices:

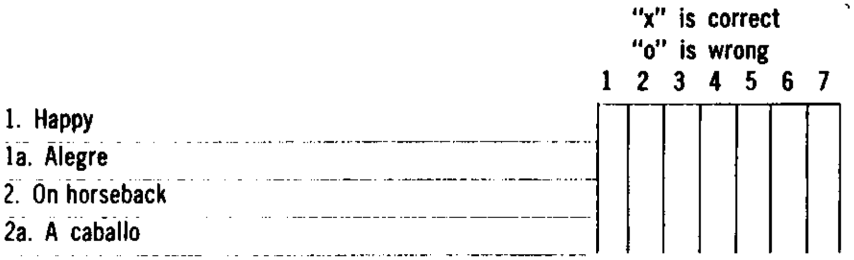
1. Unit presentation of material to an individual student, usually in the form of a question or other stimulus requiring a response such as a word to be translated
2. Provision for student response
3. Feedback or knowledge of results as to the correctness of response.

In this study, Spanish vocabulary was taught to ninth-grade English-speaking students using a teaching-machine device which could be programmed to handle either multiple-choice items or items requiring the construction of an answer. More specifically, the device used was simply a cardboard file folder with a window cut out which permitted the viewing of only one line at a time. The programmed material was presented on mimeographed sheets placed inside the folder. The first line on the mimeographed sheet contained the stimulus item—in the case of the constructed-response group, an English word—and the following line contained the correct Spanish translation. The student looked at the first line, responded on a separate sheet of paper by writing the Spanish word, moved the mimeographed sheet up one line, and compared his response with the correct answer. If his response was correct, he marked an "x" in the exposed margin of the mimeographed sheet; and if his

---

This paper is an edited summary of Dr. Fry's unpublished doctoral dissertation at the University of Southern California, 1960. Dr. Fry, Assistant Professor of Education at Loyola University, Los Angeles, is a consultant on the staff of the Electronics Personnel Research Group, at the University of Southern California.

response was incorrect, he marked an "o" in the exposed margin of the mimeographed sheet. This device was first reported by Ferster and Sapon (1958) in an experimental demonstration of teaching German to students at Harvard.

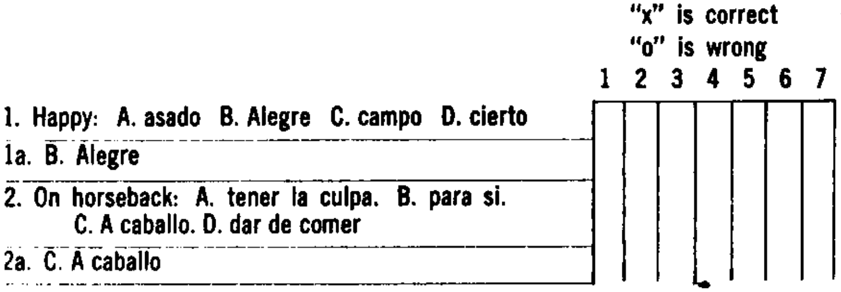


"x" is correct
"o" is wrong
1 2 3 4 5 6 7

1. Happy
1a. Alegre
2. On horseback
2a. A caballo

FIG. I. An example of the constructed items appearing on the mimeographed sheet inserted in the window folder.

Students who formed the multiple-choice response groups used exactly the same type of folders, except that the stimulus items on the mimeographed sheet consisted of a multiple-choice item. See Figure II.



"x" is correct
"o" is wrong
1 2 3 4 5 6 7

1. Happy: A. asado B. Alegre C. campo D. cierto
1a. B. Alegre
2. On horseback: A. tener la culpa. B. para si. C. A caballo. D. dar de comer
2a. C. A caballo

FIG. II. An example of the multiple-choice items appearing on the mimeographed sheet inserted in the window folder.

Spanish words and phrases taught were selected from the end of the regular Spanish text used in class, if the teachers reported that they had not been previously taught in class. A post-test was designed that used half multiple-choice items and half constructed items.

Just prior to training, all students were given one guidance trial in which the teacher read off the new Spanish word once and its corresponding English translation. Students were permitted to look at the mimeographed sheets during the guidance trial.

Both experimental groups were selected from within the same class by simply shuffling the two types of folders, constructed and multiple choice, before distributing them to the class. Separate directions were printed on each folder and discussed in class. Thus it was assumed that intelligence, motivation, socioeconomic, and other factors would be randomized. This method also gave the same time of day, teacher, and other such factors to both groups.

A pilot study of approximately 50 pupils who worked to a criterion of mastery (two correct responses for each item) showed that the constructed-response group learned better but also took a longer time. Thus it was determined to attempt to control the time variable. The pilot study also showed that both groups received near maximum scores on the multiple-choice section of the post-test, so the main training part of the study was made to contain more material and more difficult post-tests.

CONDITIONS USED

The major section of this study contained three conditions aimed at controlling time and repetition variables found in the pilot study.

Condition-I groups (multiple choice and constructed response), worked to a criterion of mastery; i.e., each student continued working until he had made two correct responses to each of the 16 items.

Condition-II groups were given the same instructions as Condition-I groups, except that they were both stopped at a set time prior to completion by the fastest student. Thus equal total working time was assured for all students. Condition-II students were a randomly selected section of the same classroom as that of Condition-I students.

Condition III was aimed at controlling both working time and number of repetitions. Windowed folders were not used, but, instead, stimulus items were presented to both response groups simultaneously by means of a large flash card. Multiple-choice training students responded by selecting the appropriate answer from four choices on a mimeographed sheet. Constructed-response-trained students simply wrote the translation on an answer sheet. Both groups received knowledge of results simultaneously after a time was allowed for responding by being shown the back of the flash card which contained the correct translation. As in Conditions I and II, both groups received one guidance trial in which the teacher read the 16 new words once. The students were the same students who were used for Conditions I and II two weeks earlier. Constructed-response and multiple-choice-response groups were selected within the same classroom at the same time by alternating rows.

A post-test was given immediately after training by all conditions. This post-test consisted of all 16 training items, eight of which were constructed items requiring the student to write the answer and eight of which were new multiple-choice items. A delayed post-test was given two days later with previously tested constructed item words appearing in multiple-choice items, and vice versa. The forms of the tests used on an immediate and delayed basis were rotated on an A-B, B-A basis among classroom groups (a total of seven classroom groups were used).

DISCUSSION OF CONDITIONS

Condition I is most like the usual teaching-machine training situation. The students work toward a criterion of mastery, and total working time and number of repetitions are allowed to vary with the requirements of the individual student.

Condition II controls the total working time, which is unlike the usual training situation but gives more scientific rigor toward answering the questions of which type of response is more efficient. If, for example, one method takes longer and is the more effective (as happened), we do not know where we are, because—other things being equal—the longer the training time the more learning we can expect.

Condition III abandons teaching machine principles to experimental design in an effort to maintain only one experimental variable. Gone is the individual's freedom to work at his own pace, for the time allotted for each item is controlled as are the number of times that he may respond to the item.

In these three conditions we see the educational psychologist's dilemma—to experiment with a situation which is similar to a usual training situation or to carefully control a learning experiment and have the conditions unlike the training situation. Interestingly, or fortunately, the results in this study all pointed in the same direction, irrespective of conditions used; but other studies have not been so fortunate, and thus the application of educational psychology experiments has encountered great difficulty.

RESULTS

Both immediate and delayed post-tests consistently favored the constructed-response training groups under all conditions. The multiple-choice parts of all post-tests were all very near the maximum, so that all the differences between the groups were seen on the constructed sections of the post-test. Mean raw scores for all groups can be seen in Table 1.

Table 1. MEAN RAW SCORES ON POST-TESTS

	CONSTRUCTED-ITEM TRAINING GROUP						MULTIPLE-CHOICE ITEM TRAINING GROUP					
	Immediate test			Delayed test			Immediate test			Delayed test		
	"a"	"b"	Total	"a"	"b"	Total	"a"	"b"	Total	"a"	"b"	Total
Condition I (Mastery of subject) N = 81	6.89	7.89	14.79	5.87	7.84	13.71	4.37	7.74	12.12	4.02	7.93	11.95
Condition II (Equal time) N = 66	6.35	7.73	14.29	4.84	7.77	12.61	4.23	7.80	12.03	3.43	7.80	11.23
Condition III (Equal time and repetitions) N = 153	4.48	7.83	12.41	3.81	7.62	11.43	3.50	7.91	11.41	2.12	7.71	9.83
TOTALS	5.91	7.88	13.74	4.84	7.74	12.58	4.03	7.82	11.85	3.09	7.81	11.00

NOTE: Subtest "a" used constructed items and subtest "b" used multiple-choice items.

The Chi-Square Median Test of significance was used in all cases because the data were not distributed normally (Guilford, 1956). These results may be seen in Table 2. Generally speaking, a significant difference was found favoring the constructed-response training groups under all conditions when constructed post-test items were compared.

Table 2. CHI-SQUARE MEDIAN TEST OF SIGNIFICANCE BETWEEN CONSTRUCTED ITEM AND MULTIPLE-CHOICE ITEM TRAINING GROUPS BASED ON POST-TESTS

	IMMEDIATE TEST		DELAYED TEST	
	Constructed subtest	Multiple-choice subtest	Constructed subtest	Multiple-choice subtest
Condition I Mastery	$x^2 = 18.6$ $P < .01$	not sig.	$x^2 = 8.8$ $P < .01$	not sig.
Condition II Equal time	$x^2 = 9.3$ $P < .01$	not sig.	$x^2 = 3.69$ $P < 0.6$	not sig.
Condition III Equal time and repetitions	$x^2 = 5.5$ $P < .02$	not sig.	$x^2 = 8.86$ $P < .01$	not sig.

Total working time, which was allowed to vary in Condition I, was recorded for each student; the mean training time for the constructed-response group was 14.2 minutes while the mean training time for the multiple-choice response group was 8.3 minutes. This difference was significant at the .01 level ("t" equals 8.4).

CONCLUSIONS

It is concluded that if recall is the criterion for learning, then the teaching-machine situation which causes the student to construct his responses (Method A) is a training method superior to the teaching machine situation which causes the student merely to select the responses (multiple choice), given conditions similar to those prevailing in this experiment.

In working toward a mastery of new subject matter, the use of recall-type of items in the teaching machine will require considerably more training time than will the multiple-choice type of teaching machine. . . .

There is a possibility that multiple-choice training items are more efficient if recognition is the only criteria for learning. This statement is based upon the high multiple-choice scores of the post-test for this group combined with the significantly shorter training time required to reach mastery, but further experimentation is needed.

On the Relevance of Laboratory Investigation of Learning to Self-Instructional Programming

THOMAS F. GILBERT

My introductory argument is that the science of human learning in the 50 years since the death of Ebbinghaus has provided us a net value amounting to very little; indeed, it may have saddled us with a methodological bias so hampering to our progress as to make the few valuable results hardly worth the price. In short, I shall argue from the premise that this science has been largely irrelevant to both educational engineering and theories of learning.

Examine the task of teaching 12 paired-associates. Uncritical assumption that the literature of human-learning science has any important measure of relevance to efficient means of teaching this material would lead us to pay careful attention to such variables as distribution of practice, speed of presentation, and the like. This would be a waste of time. I could describe in simple, systematic, and generalizable terms how these paired-associates can be taught "in one trial" [see Gilbert 1958 b, p. 632. —Eds.] to practically any unpracticed student with assurances of perfect recall after six months, a year, and even longer. I don't suppose this result should startle psychologists, although I have chosen the example simply because many colleagues find it incredible. If I assumed that we have used the classical human laboratory to discover efficient conditions for learning, I would have to conclude that my results will produce a major revolution in the science of memorization.

Another example will illustrate some of the problems involved in applying the prevalent methodology to the programming of education. Suppose we question Skinner's contention that a student should construct an answer to a problem in a teaching program rather than respond to a set of multiple-choice alterna-

---

This paper was presented at a symposium at the 1959 convention of the American Psychological Association at Cincinnati, Ohio. Dr. Gilbert, who has done extensive work in programmed learning at Harvard University and at the Bell Telephone Laboratories, is currently at the University of Alabama.

tives. We may choose to put the question to experimental test. After all, the human-learning literature gives us very little to go on. According to my estimates, this experiment has been done, or soon will be done, by 20 different investigators. I judge it to be the most popular experiment current in education. However, since I know what the experimental results are, I shall describe them, hoping to save others time.

Briefly, the experiments comparing multiple-choice selection with answer construction will demonstrate considerable individual differences on whatever criteria are used. Some investigators will, with statistical equivocation, conclude in favor of the construction method; some in favor of multiple choice; some will fail to conclude anything; and some will qualify their findings with reports of statistical interactions.

There are reasons for the individual differences. Let us examine the problem, for it is instructive. Suppose *all* experiments carried on in the statistical tradition concluded in favor of the multiple-choice method. Does this tell us that it is preferable for the student to recognize the answer rather than construct it? Regard the student who has read a problem and has before him a set of multiple-choice answers. Any number of possible behaviors are open to him. He might, for example, read the first alternative very carefully and judge it right or wrong, carefully repeating this until he found one he judged correct. As a variation, he may continue reading the alternatives even after he has found one he thinks correct, giving himself the option of changing his judgment. Again, he may read all the alternatives carefully one or more times, reserving his judgment until later. Or, he may rapidly glance back and forth over the alternatives looking for an easy cue to the correct one. Or, he may make a wild guess. Finally, and of particular importance here, he may read the problem, actually construct an answer, and then look at the multiple-choice alternatives for one which matches the answer he constructed.

With these possible behaviors in mind, let us re-examine the results of our hypothetical experiments. If an experiment concluded in favor of the so-called recognition method, what would it tell us? The so-called recognition method permits many ways of recognizing but also permits construction and guessing. To call the recognition procedure a "method" is to travesty the very word. If the results favored multiple choice, it may be only because all students in the multiple-choice group first constructed an answer and were forced to change it when they found no match. If this were truly the case, we would conclude erroneously in favor of recognition and, more importantly, obscure the opportunity to discover a better way to use answer construction. This by no means exhausts our sources of error. Suppose no one in the multiple-choice groups constructed an answer before looking at the choices. We still could not say that the construction method was inferior simply because we obtained examples of only one form of the method. The scientist who reports such results, or acts upon them as authoritative, has simply evaded his responsibility for description. The science is charged with reducing complex events to a manageable set of variables, but it cannot do this while unaware of the very variables with which it has to work.

I am less concerned with disclosing the irrelevance of the methodology and data of the classical human-learning laboratory than I am with prescribing a remedy. Now may be the time when prescriptions are sorely needed. This new endeavor to improve education, recently stimulated by B. F. Skinner, shows signs of growing up like Jimson weed. For evidence of uncritical and precipitate development I have to go no further than to point out that there probably exist 100 different mechanical gadgets oddly called "teaching machines" and, in sharp contrast, maybe no more than two or three teaching programs which in any way could be called complete. Of deeper significance, I think, is the fact that there is a whole rash of so-called "control-experimental group" experiments purporting to answer questions about principles of programming education. Since I believe that there is a basis for a more considered effort, I should like to describe two alternative sets of methodological rules. The first set of rules may be useful to anyone whose foremost concern is the improvement of a given educational subject matter. The second methodology will apply to developing more general rules for programming education independently of a specific subject matter.

My qualifications for the job of providing the first set of rules are quite meager and are based on little more than a three-year-old avocation. My most reliable claim to authority in this matter is the fact that few men have seized more opportunities to make mistakes in this field than I have. These rules for getting done the immediate job of programming a specific subject matter I offer with no particular pride, but with some confidence and in dead seriousness.

RULE 1. If you don't have a gadget called a "teaching machine," don't get one. Don't buy one; don't borrow one; don't steal one. If you have such a gadget, get rid of it. Don't give it away, for someone else might use it.

This is a most practical rule, based on empirical facts from considerable observation. *If you begin with a device of any kind, you will try to develop the teaching program to fit that device.* The so-called "teaching machine" is a disease, not a challenge to self-control, and the only safe cure is to get rid of it. The recommended treatment is the cold-turkey method—don't try to taper off on programmed or scrambled textbooks.

RULE 2. Resist the temptation to design formal experiments. You don't want to know *whether* one method teaches better than another, you want to know *what* method teaches best.

This rule is based on the simple logic that a really efficient method of teaching a thing would display itself in a "control-experimental groups" study only if a really efficient method were used in the experiment. Thus you could conclude from the experiment that a method was unusually efficient only if you already knew it. In short, I am saying that the first function of this teaching laboratory is as a place of discovery, not a place to prove preconceptions.

RULE 3. Your prime purpose is to provide a student with a behavior repertory called a subject matter. If that behavior repertory is, say, physics, your problem is to take him there from whatever repertory he now has which even vaguely approximates physics.

This rule is stated to emphasize the fact that a subject matter is a class of behaviors and that everyone has some behavior which approximates that behavior class. It is easy to forget that the behaviors one goes through to master the subject matter may be different from the actual subject matter behaviors. The failure to grasp fully the implications of this rule has been, in my experience, the biggest single stumbling block for people learning to program education. The natural tendency is to begin by breaking the subject matter down into small, concise units. While this is valuable for describing the repertory you wish to build, these behavior units usually are not the ones which will actually build that repertory. They are test items, not teaching guides.

RULE 4. Get yourself an expert teacher of the subject matter you wish to program. Be wary of a college professor; he may never have seen a student learn. Remember that a good teacher is a more complicated, flexible "teaching machine" than you could possibly build. If you can't get a good program into him, you will never get one into a mechanical gadget.

This rule is not meant to suggest that the teacher is to tell you how to build the repertory. Quite to the contrary. The student will tell us this. The teacher is only the place to start.

RULE 5. Get yourself one student. I repeat, *one* student. You are about to perform an experiment in which you are permitted no degrees of freedom—that is, if the word "self" in "self-instructional" can be taken seriously. Once you have discovered an efficient program for one student, you will have described the gross anatomy of the most generally useful program.

RULE 6. You have to start somewhere, but forget that you are an expert on human learning. You aren't. Assuming you are the teacher, you should begin with the most trustworthy facilities you have available: First, trust your common sense; next, use the approximations to principles of programming that have been set down by a few people. Remember, these people probably are not more expert than you, only more audacious. They may be mostly wrong. Use their principles only as a starting place.

RULE 7. Obtain the following materials: paper, pencils, and index cards. Use no gadgets unless they are part of the subject matter. For example, if you are programming home economics, you may need an electric toaster. Resist the temptation to use the toaster as a "teaching machine."

RULE 8. You are now ready to begin programming. Think of the process as an exploratory experiment in which you do not know what the effective variables are. Your problem is to discover them. Using index cards, write out a series of questions, probes, etc., to which the student can respond. Write these items in a way that you think will lead him to a mastery of a small part of the subject matter.

It seems a good idea to write these items while having an imaginary student before you with whom you are carrying out an imaginary interchange.

RULE 9. Take your first crude effort to the student. Remember, he is going to teach you. This student cannot fail. If he doesn't get where you want him to go, you have failed. Try something else. In the absence of anything better, let whim be your guide. If you come to a dead end, vary your approach until you have gotten him where you want him to go. Tape record all sessions. The important thing to remember is to keep varying your behavior until you are successful and to describe what you do.

RULE 10. Once you have learned how to get the student through part of the material, keep going. Pay the teacher and student so they won't leave you. This can be dull work. Don't invest much time in constructing materials before you have tried them out on the student.

RULE 11. Once the teacher really appreciates immediate reinforcement and has discovered that the student alone can tell him how to teach and once he has learned how to keep varying his approach, he is more of an expert than you. Remember, it is easier to teach a physics teacher what you know about programming than it is for him to teach you physics.

RULE 12. Take your time. Education has been waiting for you since the dawn of history. When the student has the repertory you wanted to build, and when you can describe how he got that repertory, you are ready for the next step. Edit the material and try it on another student. Make whatever changes necessary for your program to take care of both students. After fewer than 10 tries, you will have a program which will teach 98 percent of the students. And you will have discovered how to adjust the program for individual differences.

RULE 13. Don't be too concerned that your program is not perfect. It works. It can always be revised. If you have followed the rules . . . you have done a respectable job. You can describe an exact procedure for getting most students from oblivion to mastery—not to a C average. Now, and only now, you are ready to think about automation.

RULE 14. Prepare to automate the program by discarding any "teaching machines" which, in your weakness, you kept around. Remember that you have a teacher who is a vastly complex machine and you have discovered how to make him work with efficiency. All you need to do now is to substitute more economical devices for the teacher's operations wherever you can. You probably will end with several devices. Examine each operation and fit a device to it. Never let the device dictate the program.

Suppose you have programmed physics. In examining how the teacher builds a masterful repertory in these students, you may see that he came to teach the mechanics by verbal discourse which you could simulate on a device similar to Skinner's "teaching machine" or on a paper moved under a cardboard mask. Perhaps the only possible substitute for the teacher's necessary demonstrations of certain electrical phenomena is a motion picture film or television. The compromise between efficiency and economy is your guide. There is no reason to build a machine to do what a cardboard mask or programmed-text would do as well. If it becomes evident that there are things which can be done only by a human, without resorting to excessive expense, then you simply have to write an exacting program for that human. If the program calls for an inspiring lecture, and the expense is justified, include such an inspiring lecturer amongst your instruments. After all, we are programming the subject matter, not fitting the subject matter to some preconceived machine or method. It is an idle question whether "teaching machines" or television can teach. The real and answerable question is, "How can we teach?" and the rest is a matter of classroom economics.

These rules are designed to guide one who wishes to apply whatever we already know to the task of engineering a specific educational problem. If we wish to discover principles which are not bound to a particular problem, and are still useful to the problems of education, we must resort to a much more expensive and elegant laboratory. There is no middle ground.

This more elegant laboratory we will approach with a very similar logic if we wish to provide a generalizable and relevant set of human-learning principles. The logic I refer to is that of operant conditioning methodology, and it has been described elsewhere, notably by B. F. Skinner. It deserves restatement in the context of the study of human learning.

The first characteristic of the human operant conditioning laboratory is the provision for intensive, long-term study of a single human in an artificial and fully controlled environment. This logic is based on the universal agreement that the human comes to the laboratory with a complex conditioning history. The traditional approach has attempted to get behind idiosyncracy by statistical averages of short behavior samples. What this gives us instead of generalizable and timeless principles is an average of cultural effects. The traditional laboratory is too artificial for an interesting sociology, too sociologically sensitive for a culture-free (diachronic) behaviorism. Students may prove to use the recognition "method" better on the average simply because they have been trained that way, on the average. The diachronic laboratory, I am suggesting, will have subjects spend a great part of each day for many months in an artificial environment, hoping that environment and its associated consequences will eventually take on a unique character which will override many of the effects of conditioning in other situations. In short, we wish to make this laboratory a unique and standard culture.

Secondly, we use no, or minimal, instructions. Instructions are used as a means of getting the experiment under better control. They serve as extremely complex reinforcement conditions —the subject follows them because he fears the instructor, or wants to impress him, or feels kindly toward him; who knows? They are compounded out of an intricate sociological past. Instructions bias the experiment by telling the subject what to do. We conceive of that subject as a part of nature who is never "wrong"—and who is to instruct us in his ways. Instructions would be standard only if they had describable and similar effects on every subject.

For purposes of control we use precisely manipulable consequences, the reinforcement value of which can be demonstrated . . . nickels, food, etc. Our subjects earn about three dollars an hour; and if they don't quit when the money is removed, they may be working to make fools of us. Instead of instructing a subject to behave in a certain way, we make his nickels contingent upon that behavior. The experimenter stays out of the experimental room. We don't know nearly enough to describe the effects of his presence.

We make each subject his own control, and we can do this by virtue of the long-term investigations. We collect months of base line data while standardizing and making unique the experimental space. A single subject may work in this laboratory for five or more years.

While the logic of this laboratory requires it to be an artificial environment, the environment is not arbitrarily constructed. Every event in it has its own logic, the primary requirement being formal representativeness. The tasks of the traditional laboratory are chosen for traditional reasons, the nonsense syllables because Ebbinghaus used them. There are many tasks which represent serial behavior functions far freer from cultural associations than these syllables. In the diachronic laboratory we do not wish to represent specific cultural forms. A truly diachronic laboratory will allow us to obtain data from Fiji Islanders which would have the same meaning as data obtained from New Yorkers. Artificial tasks with near zero past associations are impractical unless you can work with your subjects for many hours. The education of a single human takes many years. The scientific knowledge relevant to this education will not be gained by one-hour sessions with subjects in a laboratory hardly more rigorously contrived than the classroom itself.

The most important turnabout in the traditional laboratory logic is the conception of the very purpose of that laboratory. It is usually conceived as a place where we put to test the effects of variables chosen on a priori grounds. The a priori choice is sometimes dictated by a theory; more often, by traditional and economical considerations. What we have called "experimental design" is more accurately designated as "pre-experimental design." This a priori selection of variables so biases our conception of nature that we sometimes throw out a subject who doesn't behave in accordance with preconception or statistical requirement. The rare subject who learns the paired-associates in one trial is the fellow who gets tossed out, either because he threw no light on some mediational hypothesis or because he violated the normal distribution. I would suggest that we would learn more if we threw out all subjects except this one.

This conception of the laboratory as first of all a testing ground biases the experimenter to ask such questions as: "Is this variable more effective than that one?" or, "Can television teach?" As a result, the laboratory is not seen as a place for discovery. It is the discovery role which will lead us to emphasize such questions as, "What variable is effective?" and "What can teach?" —these being the questions relevant to education. Under such logic, "experimental design" becomes the design of a laboratory which will allow us to get all possible "recognition" behaviors overtly expressed.

I can best demonstrate what we stand to lose when we deemphasize the discovery function of the laboratory by relating to you an experience of a friend. This friend tells the story of his last venture into the laboratory investigation of human learning when some years ago, as a newly produced experimental psychologist, he set about to study human avoidance conditioning, presumably because there was an old hand-shock grid cluttering up the laboratory. He was well equipped in the classical tradition. He had a theory (something about an inscrutable mediating between light and shock). He had a group of male college sophomores, carefully matched for sex and education. He had an experimental design, a properly conservative two-tailed *t* test. And to complete the Galilean picture, he had a standardized source of uncontrolled variation: a wholly mysterious set of instructions designed to startle the most self-satisfied mediator out of its hypothetical existence. The first of the 60 scheduled sophomores proved too unsophisticated for such a high-level scientific setting. When the experimenter threw the switch which flashed a light and pulsed a shock, the subject blandly kept his hand on the grid. Trial after trial the shock intensity was increased without a twitch from the indifferent hand when, on the final trial, with the current at a searing maximum, the subject's steady fingers began to smoke. At this point, the experimenter's fine consideration for scientific control gave way to understandable human curiosity, and he asked the subject why he didn't move his hand. The subject's reply provided perhaps the finest opportunity for scientific discovery since Skinner's first automatic food-magazine broke down. The subject simply said, "I thought you were studying how much *guts* I had!" My friend thereupon cancelled the schedules of the remaining 59 sophomores, and today he is productively engaged in the investigation of the eating and defecating behavior of hooded rats. Alas, he may have foregone the opportunity to provide us the first systematic and quantitative methodology for the study of human courage.

I relay this story in all seriousness because it served to illustrate my thesis. In this not unreasonably overdrawn example, the investigator may have failed to reduce guts to a science, but he thankfully learned in a hurry a fundamental which it is taking, or has taken, many of us much longer to learn. That fundamental is that we are not going to increase significantly our already considerable market-place knowledge of human behavior with the pedestrian methodology prevelant today. This methodology was unsystematically pieced together by the sheer inertia of tradition, the peculiar economics of college professing, and an indecisive eclecticism.

Problems In Programming Verbal Learning Sequences

LLOYD E. HOMME and ROBERT GLASER

IT HAS been over five years now since Skinner originally introduced the notion of using a machine to teach. At that time many of us here had an overwhelming conviction that we were on the threshold of a revolution in education. Thus far, we must sadly admit, the revolution has not materialized. If the metaphor is at all apt, we may say that we are still standing on the threshold, although we may have a toe or so over it. Most of the teaching this year will be done in pretty much the same old way it was five years ago. (Notable exceptions are Earlham College, Hamilton College, and Harvard University.) Teachers will typically stand in front of the class and lecture and write on the blackboard and periodically give tests. Why do we continue in this fashion? We submit to you that it is not timidity or caution which is holding us back. The reason most courses will be taught the same old way this Fall is *not* that we are awaiting data concerning "size of step" or optimal cueing procedures, nor that we need to debate whether the process of learning by machine should be called "shaping" or "discrimination training." The difficulty is *not* that we do not have teaching machines. As Thomas F. Gilbert has pointed out, there must be dozens of teaching machines of various kinds in this country today. Everyone in this room will agree upon the source of the difficulty. The simple truth is that the reason most teachers will not be using or trying out automated teaching is that their courses are not programmed. It is for these reasons that we are devoting this paper exclusively to the problem of programming.

How did we get into this ridiculous situation in which machines outnumber programs literally by a large factor? We do not have far to search for the reasons. Programming is an aversive task which can be successfully avoided by designing and building teaching machines. Why is programming so aversive? The reason is that programming is time-consuming and difficult. At least part of the trouble lies in the fact that there are no rules for programming of sufficient precision to be helpful. The beginning programmer is told, for example, "to introduce new concepts carefully." He has the perfect right to ask: "What do you mean by *introduce*? What do you mean by *carefully*?" At the present time only vague, unsatisfactory answers to these questions can be given.

The programmer actually needs two systems: (a) a set of rules for analyzing knowledge to be installed in the subject and (b) a system for frame-writing to get this knowledge into the subject's muscles. Now, we will be concerned only with the second category, that of frame-writing. We will assume that the subject matter has been analyzed at some prior time; that is, we will assume that the programmer knows what behavior he wishes to establish.

It is easy to specify some of the properties a system for frame-writing ought to have. The system ought to be general enough so that it is germane to any subject matter. The system ought to be specific enough so that, given the pertinent information, the construction of a particular frame or series of frames should become almost a mechanical matter. Ideally, what we would like is a system with sufficient precision so that information about a subject matter could be fed into a computer with the expectation that a completed program would result. An approximation of this kind of precision is to attempt to instruct a subject-matter expert how to write frames and construct a program in his own subject.

In searching for systematic rules for programming, we might take seriously, first of all, that the program with which the student comes into contact acts like a tutor. This leads us to examine carefully the verbal behavior of a tutor or a tutor-pupil interaction. The most impressive thing about this tutor-pupil interaction is its complexity and the apparent difficulty involved in analysis. Not only is one dealing with the most complicated organisms known, but with the most complicated behavior—verbal behavior—of which these complicated organisms are capable. Further, these two organisms, the tutor and student, are interacting with each other and possibly with the verbal behavior of still another organism in the form of a text. In the face of such overwhelming complexity, it is perhaps not too

---

This paper was presented at a symposium at the 1959 convention of the American Psychological Association.

surprising that rules or principles specifying how the teacher should proceed in this situation are almost totally lacking. It is also clear that, since this interaction is so complicated, any abstraction, or any system which is developed, will be open to criticism on the grounds of oversimplification. We are willing to take this risk.

Despite the complexity of which we have complained, it is possible to detect a sequence of events which appears to be common to the tutor's verbal behavior whether he is tutoring in German, statistics, mathematics, or whatever the subject matter. This is that a tutor will typically enunciate a principle and follow this with a series of instances of the principles. The statistics tutor, for example, is likely to say: "The mean is simply defined as ΣX over N. For example, in the distribution below . . . ." The foreign language teacher might say: "Such and such takes the objective case. For example, . . ." The psychology professor says: "The discriminative stimulus is a stimulus which sets the occasion for a reinforced response. For example, if we have a light in a Skinner box . . . ." An almost endless series of examples like this could be cited, but the point is clear: In any tutorial episode a large part, at least, of the tutor's verbal behavior may be easily classified into two categories: (a) principles or abstractions and (b) instances or special cases of these abstractions. For the sake of simplicity, let us speak of these from now on as "rules" and "examples."

The question we wish to raise now is how much of the verbal repertoire called knowledge of a subject matter is encompassed by these two classes of a verbal behavior, rules and examples? For an answer to this question, we turn now to a band of intellectuals frequently ignored by psychologists, the logicians. Philosophers of science say that knowledge grows somewhat in the following fashion. A scientist notices the recurrence of an event frequently enough to be interested in it and gives it a name. This is "naming," the lowest level of law or abstraction. Once the circumstances under which an event occurs are identified, a higher level of abstraction, a law, is said to have been generated. If a number of laws can be subsumed or deduced from a broader or more inclusive concept, this constitutes a theoretical statement, or a lawful relationship of a still higher order, and so on. Now, the logicians are interested in discriminating between these various kinds of statements, such as discriminating between merely naming an event and stating a lawful relationship concerning the event, or discriminating between kinds of laws and kinds of statements—true, false, transitive, and so forth. But, since we are interested in oversimplifying the world, we may, rather than ask the question, "How do these event classes differ?" ask the question, "In what respects are all these statements alike?" We wish to assert that this population of statements which we have described can be *exhaustively* analyzed into rules and examples of varying degrees of complexity and abstractness. We are reminded also that whether any given statement in this population is classified as a rule or an example is a function of the level of discourse. Any given statement is a rule for the instances which lie below it, but at the same time is an instance of the rules which encompass it. We argue that this sort of a simplification into rules and examples regardless of the level of abstraction is a defensible one; and since the same laws of learning apply to the learning of statements of a low order of complexity as of high, we need not discriminate among them.

What is the psychological status of rules and examples? It is to be noted that rules and examples may function as extremely powerful prompts. A rule is almost the archetype of the thematic prompt; it may share none of the formal properties of the behavior it is intended to evoke, yet makes the behavior extremely probable. Rules can also be formal prompts, of course, and if they are, we are likely to call them mnemonic devices. Examples also, of course, can vary in thematic and formal prompting properties. When a tutor or textbook cites a rule and follows it with a series of one or more examples, the hope is that the rule will become sufficiently strengthened so that it may serve as a thematic prompt when on future occasions the student or reader encounters similar examples. The teacher, whether he will admit it or not, is engaged in altering certain conditions of stimulus control of behavior. If this is true, then we should expect to find the rule-example sequence in places other than the classroom or textbook. And it is, of course, true that the rule-example sequence is extremely prevalent wherever stimulus control is to be altered. The rule-example sequence seems particularly noticeable when persuasion is involved, qualifications to a statement are being specified, or instructions are being given. These conditions do not exclude everyday speech, of course, and the rule-example sequence can be detected here also. Rule: "The Pirates are a bunch of bums!" Example: "Look at the way they played Saturday!" Since the rule-example pattern is so prevalent, we are constrained to inquire about where it does *not* occur in the verbal world. We frequently do not find the rule-example pattern in news items or in technical articles. In both cases, apparently the stimulus control is deemed to be adequate at the beginning: in the case of the news item, because the relationships are extremely simple; and in the case of the technical article, because of extensive training on the part of the reader. Another place one does not find the pattern is in rote learning, simply because there are no rules in this game. What the subject has to do in this case is to make up his own rules, or prompts, as Dr. Gilbert has pointed out. Still another case in which the rule-example pattern is not followed is in the usual concept formation experiment. Here, too, the subject is simply presented with what we would class as examples and is required to deduce the rule.

We need only to add one more small step now in order to define, crude as it is, the *ruleg* programming system. We have already divided the verbal world into two kinds of statements, rules and examples, which we will abbreviate as follows: *ru*s for rules and *eg*s for examples. Now we note that, since programs typically are comprised of complete as well as incomplete statements, we can further divide *ru*s and *eg*s into two categories, complete and incomplete. We have, then, four categories, complete and incomplete *ru*s, complete and incomplete *eg*s. We abbreviate incompleteness by placing a tilde (∼) over the *ru*s, or *eg*s: $\widetilde{ru}$ and $\widetilde{eg}$. Two more categories complete the system: false *ru*s and false *eg*s. These we abbreviate simply with the negation sign, a bar over the *ru* or *eg* yielding $\overline{ru}$ and $\overline{eg}$. What we call the *ruleg* system is simply a name for programming by means of combining statements of these categories in such a manner as to comprise a program of the machine-teaching type.

Our general strategy of programming is to attempt to take seriously the contention that a program is primarily an instrument for establishing discriminative control over behavior. Further, we assume that the response the S makes is the one

which receives the greatest amount of strengthening. As a result, we see that our first job is to get the discriminative operant emitted in the presence of the discriminative stimulus. Table 1 of the present paper is a modification of the very clever and excellent illustrative program prescribed in Table 2 of Skinner's article in *Science* (1958). The program prescribed here in Table 1 is virtually the same program, reconstructed according to the principles we have just been discussing. Turning to Table 1, Frames 1 and 2, we may summarize the *ruleg* system briefly

Table 1. SKINNER'S HIGH-SCHOOL PHYSICS PROGRAM RECONSTRUCTED ACCORDING TO THE RULEG SYSTEM

CLASS	SENTENCE TO BE COMPLETED	WORDS TO BE SUPPLIED
$ru + \widetilde{eg}$	1. To "emit" light means to "send out" light. For example, the sun, a fluorescent tube, and a bonfire have in common that they all send out or .... light.	emit
$\widetilde{eg}$	2. A firefly and an electric light bulb are alike in that they both send out or .... light.	emit
$ru + \widetilde{eg}$	3. Any object which gives off light because it is hot is called an **incandescent** light source. Thus a candle flame and the sun are alike in that they both are .... sources of light.	incandescent
$\widetilde{eg}$	4. When a blacksmith heats a bar of iron until it glows and emits light, the iron bar has become a(n) .... source of light.	incandescent
$\overline{eg}$	5. A neon tube emits light but remains cool. Unlike the ordinary electric light bulb, then, it is not an .... of light.	incandescent source
$\widetilde{ru}$	6. An object is called incandescent when	it emits light because it is hot.
$ru + \widetilde{ru}$	7. It has been found that an object, an iron bar, for example, will emit light if its temperature is raised above 800 degrees Celsius. Therefore we say that above .... (temperature) objects will become	800° Celsius incandescent
$ru + \widetilde{ru}$	8. An electric light bulb produces light when the fine wire, technically called a **filament**, inside the glass is heated to incandescence. This means, then, that the fine wire or .... must exceed a temperature of about ....° Celsius to emit light.	filament 800
$\widetilde{eg}$	9. In an electric light bulb an electric current is passed through the fine wire or ...., it becomes ...., because it is heated to a temperature above 800° Celsius.	filament incandescent

Table 1. SKINNER'S HIGH-SCHOOL PHYSICS PROGRAM RECONSTRUCTED ACCORDING TO THE RULEG SYSTEM (continued)

CLASS	SENTENCE TO BE COMPLETED	WORDS TO BE SUPPLIED
ru + eg + $\widetilde{ru}$	10. The hotter an incandescent light source becomes, the greater the amount of light it emits. For example, an object heated to 900° Celsius would emit more light than the same object heated to 800° Celsius because the .... of light emitted depends on the .... of the object.	amount temperature
$\widetilde{eg}$	11. A nearly "dead" battery in a flashlight produces only a feeble light in the bulb, because the filament fails to get .... enough to produce much light.	hot heated
$\widetilde{ru}$	12. An incandescent light source can be made nonincandescent by cooling it, since the amount of light emitted depends on the .... of the source.	temperature
ru + $\widetilde{eg}$	13. The **color** of the light an incandescent object emits changes from **red** to **white** as the object's **temperature** increases. Thus, when a blacksmith heats a piece of iron it first glows .... (color), and then as it gets hotter turns	red white
$\widetilde{eg}$	14. A weak flashlight battery produces only a dull .... (color) glow in the bulb's filament and little light is emitted because the .... of the filament is relatively low.	red temperature
$\widetilde{ru}$	15. Both the .... and .... of light emitted by an incandescent object depend on the temperature.	amount color
ru + $\widetilde{eg}$	16. In "emitting light," an object changes or, technically, "converts" one kind of energy to another. For example, in a flashlight the electrical .... supplied by the battery is changed or ....ed into heat and light.	energy converted
$\widetilde{eg}$	17. The light from a candle flame comes from the .... released by chemical changes as the candle burns.	energy
$\widetilde{eg}$	18. One "turns on" an incandescent light bulb by closing a switch so that .... energy can be converted to .... and	electrical heat, light
$\widetilde{eg}$	19. The hot wick of a lighted candle gives off small pieces or **particles** of carbon. These ....s are heated to incandescence, so that we know their temperature must exceed ....°	particles 800° Celsius
$\widetilde{eg}$	20. Any incandescent object can be made nonincandescent by cooling it below 800° Celsius. For example, if one places a piece of metal in a candle flame the carbon ....s will be .... below incandescence, and will collect on the metal as soot.	particles cooled
$\widetilde{eg}$	21. Smoke from a candle is comprised of carbon particles which did not burn and cooled below .... (temperature) after they left the heat of the flame.	800° Celsius

Table 1. SKINNER'S HIGH-SCHOOL PHYSICS PROGRAM RECONSTRUCTED ACCORDING TO THE RULEG SYSTEM (continued)

by example. The first frame is composed of a *ru* and an $\widetilde{eg}$: the *ru* being "to emit light means to send out light;" the $\widetilde{eg}$, of course, being the next statement beginning with the phrase, "For example," which the student is asked to complete. The *ru* here is a prompt for the $\widetilde{eg}$. The second frame contains just an $\widetilde{eg}$ without the prompting rule. It is worthwhile contrasting this procedure of getting the student to emit the response immediately in the presence of discriminative stimuli and then immediately getting him to emit the same response without prompting on the second frame, with Skinner's procedure to strengthen the word, "emit," by having the student read it in context several times before the programmer asks for it. Skinner uses the word, "emit," in his Frame 7 for the first time, and the student is not required to emit the word until the Frame 16. (See pages 148-49.)

Concept formation, which demands generalization within classes and discrimination between classes, is easily dealt with in a program, as both Skinner's and the *ruleg* programs illustrate. Generalization within classes is easily accomplished with a series of $\widehat{eg}$s having as large a range as is required. In the present program, for example, the $\widehat{eg}$s range from statements about tiny carbon particles to the sun. Discrimination between classes is dealt with by including $\overline{eg}$s, that is, examples which are not instances of the rule in question. In the present illustrative program, for example, the neon tube and the firefly are included to establish a discrimination between incandescent and nonincandescent classes of light sources. The $\overline{eg}$ is presented in Frame 5.

You will, of course, judge for yourself by comparing the two program segments, the relative adequacy of this quasi-formal system, but it should be remarked that our group has found this simple systematization quite helpful. First of all, it appears to result in some economy in programming; second, it appears to facilitate talking about programs—we can now speak of program properties in terms of *ru*s and *eg*s without undue confusion; third, it appears to facilitate instruction in programming to a marked degree. Our group, for example, has instructed undergraduates in programming with quite encouraging results. But, perhaps most important, it has facilitated the programming process itself. The programmer is no longer left to gaze out the window while he cogitates about a sly way of "approaching a concept." If he can state the rule and if he can state examples of the rule, he is off to a good start in a program.

The present paper has dealt with the problem of getting rules, principles, and instances from the programmer into the subject's muscles. It has not dealt with the equally important collateral problem of identifying all of the intraverbal connections which should be established. Although there is insufficient time to dwell on the topic, it might be mentioned that the use of verbal "matrices" can greatly facilitate the discrimination of intraverbal connections (IVC's) to be strengthened. Briefly, the matrix technique, invented by our colleague James Evans, is employed somewhat as follows: The important terms (or shorthand for *ru*s) of a subject matter are placed on both axes of a matrix, and each intersecting cell specifies a possible intraverbal connection. Suffice it to say at this time that through this technique large numbers of IVC's are specified from which the programmer may select for inclusion in a program.

What has all this to do with research issues in programming the topic of this symposium? Well, in the first place, it is quite impossible to do research on teaching machine techniques without programs, so that, to the extent that the programming aspect is facilitated, research may be facilitated by the development of suitable programming techniques. In the second place, there is a distinct possibility that research issues and variables to be investigated may change in importance as more is known about programming itself. For example, in our own case the problem of the fading of prompts has diminished in importance

since we have found that usually we can immediately withdraw the prompt if we have once gotten the behavior out in the presence of the SD (discriminative stimulus). Similarly, the size-of-step variable fails to concern us in the same way that it used to. We used to think about how many steps it would take us to "work up to a concept"; now, as Table 1 indicates, the problem is more closely described as, "How quickly should extensive variations in the *eg*s occur?" and, "How many *eg*s are sufficient?" The collateral problem remaining here is a systematic way for breaking *ru*s into sub-*ru*s which requires strengthening. For example, we cannot start out with the *ru:* "An object which emits light because it is 800 degrees Celsius is called incandescent." We must first break down the *ru* into sub-*ru*s, an example of which is the little series on the word, "emit." Another field for experimentation suggested is the ordering and amount of strengthening of *ru*s and *eg*s. As a further illustration of our contention that research issues may change as our competence in programming grows, we may cite one anecdotal example. One thing of which we were all sure is—or was—the importance of confirming the student's answer immediately after he had given it. Now we are not as sure as we used to be. As part of our procedure of programming, the programmer typically writes out a series of frames in longhand, then tries these frames out on a subject or two, then the frames are typed and corrected on the basis of the subject's responses. At this stage the typed, slightly neater, series of frames is tried out on additional subjects. It is this last phase which is of interest for the following reasons. More than once the cards have been handed to these subjects with the answers inadvertently omitted from the back. To our surprise, this made no difference to several subjects. At least that was their claim. They said, "Oh, it didn't bother me, I knew my answers were right anyway." Examination of the student's responses showed that this was indeed correct. This was only a fragmentary observation, of course, but there is at least the possibility that the confirming answer need not appear on every frame providing the program is a good one. We intend to look into this very shortly.

The general research strategy which we have finally adopted appears to be exactly that which Skinner has held all along; that is, to construct a complete program which does the job it set out to do—teaches the subject matter. It is at this stage that we believe that experimentation can be profitable. Once having a program constructed which teaches perfectly, or nearly so, one can then sensibly ask: "What can be done to make this program more efficient? Can anything be done to the program to shorten it and still leave it 100 percent effective, or to make it more reinforcing to the student?" and so on. There are other things one could do if one had a complete course programmed. One could find out what alterations in the technique or in the program were necessary in order to adapt it to younger age groups, to the feeble-minded, to the aged, to the doctoral candidate, and so forth. In short, we are convinced that if psychologists were armed with complete course programs, they could bring about some startling events in the world. We hope that five more years will not have to elapse before they are given this chance.

Some Perspectives and Major Problems Regarding "Teaching Machines"

SIDNEY L. PRESSEY

THE WRITER has for many years been active in development of and experimentation with "teaching machines" but is no longer; seen from this position, which should give both perspective and detachment, the current "explosion" of great interest in this topic involves (in his view) certain major problems warranting brief comment.

PSYCHO-EDUCATION PROBLEMS

Several admirable reviews have appeared of relevant laboratory experimentation and theories of learning derived therefrom, but there has been considerable neglect of the analogously relevant research and theory in education and even in educational psychology. In apparent defense of that neglect, the flat statement has been made (Galanter, 1959, p. 14) that "educators have not done the systematic, controlled type of study that is needed to reveal general principles of learning efficiency." The writer has questioned that conclusion (Pressey, 1959, p. 187). Hilgard (1960) has recognized both the inadequacy of present theory and its possible gain from applications as in educational automation. But anyhow, if the learning theorist plans to go into business in the neighboring country of education-land, he should make a real and friendly effort to understand the language and local culture, and study closely the goods or processes he hopes to improve or supplant. His plans may thereby be facili-

---

This recent paper by Professor Pressey has not been previously published, though a modified version of it was presented at a symposium sponsored by the Department of Educational Psychology at the University of Texas in May 1960.

tated. Thus he may find educational tasks he is about to undertake have already been much delimited; for instance, word-counts have found which words are most used and so need to be spelled and error-counts have found the "spelling demons" which cause most trouble. Analyses of errors in arithmetic have suggested ways to combat or prevent errors. The theorist may even find that the exceedingly complex processes of learning which go on in the schools have characteristics which he has failed adequately to take account of in his automating, and perhaps even in his basic theories. Four issues here seem worth special mention.

Educators have long recognized the inadequacy of superficial verbal learning; so Horace Mann told of children who were glib in talk about geographical features which they could not identify though visible from the classroom window. Research has shown that knowledge of arithmetic must be grounded in countings and perceived numbers of actual things. Field trips in botany and laboratory work in chemistry are taken-for-granted features in courses in those subjects. For training in trade or profession, the value of supervised practice is generally recognized; the would-be stenographer gets some office experience, the would-be physician or clinical psychologist an internship. In the all-too-few investigations which have taken account of modes of education in relation to functioning outcomes thereof (as in safety education, agricultural education, and certain training programs in the services) the importance of frequent experience with the things and the activities involved has been emphasized. So safety education, which does reduce accidents, involves viewing of hazards in traffic (or shop) and practice in safe procedures. Like the giant Antaeus of Greek mythology, who drew his strength from mother earth and required frequent contact with earth to maintain it, effective learning must keep contact with experiential reality. Sometimes it seems presumed that programming and devices for automated education may be so adequate that such contacts are not needed, or at least need not be especially provided for. Rather, the special nature of automated education may make yet more necessary some explicit provision, in a total educational program, for such observing and doing.

It would probably be agreed that the appreciative reading of a Wordsworth sonnet involved exceedingly complex psychological

processes: sensitivities to rhyme and rhythm, evocations of images and of feelings, blendings of attitudes, some as broad as one's philosophy of life, all this and more, somehow integrated into a total aesthetic experience having an essential unity. Perhaps it would be agreed that the detailed programmer and his device could hardly intrude into the reading of such a literary gem, though he might help in elucidating some expository prose passage. But the point is that much and probably most reading (also listening and viewing as of pictures) is also enormously complex, and the values sought complex also. The detailed programmer might both exaggerate detail and hinder the larger apprehension. And his value might be questioned in a very different type of reading: the selective skimming usual in glancing over a newspaper but characteristic also of the able student's search for matter on some special topic or the scholar's continuing scanning of the literature for material bearing on his specialty. In short, the contention is that much school learning is of an integrative complexity or a selectivity (or both) beyond what automators mostly now seem to recognize.[1] Rather, their numerous little steps concern detail when detail is not the concern; and the program enmeshes, making individualized selectivity difficult if not impossible.

From kindergarten through graduate school, the most effective learning is usually, in various respects, social: in motivation and continuing stimulation from other pupils, teacher, parents, community; in guidings and elucidations from both classmates and teacher; in goal-settings and appraisals. Recent efforts at increased educational effectiveness have stressed such factors (Brandwein, 1955). Practically all the writer's experiments with automation over the past 30 years have been in informal laboratory settings with such values of sociality recognized and sought. But most recent work seems to neglect such possibilities, or even make a point that the student works by himself, the learning situation being depersonalized. In contrast, the writer would argue that use of teaching machines might better be shared fun, and could distinctively aid in further desirable socialization of school work.

---

[1] Some seem also not to realize that learning may best begin with some complexity—that it may be better to begin reading with words rather than letters. And not only a picture but also a chapter or magazine article may be glanced at for its general significance, with focus on detail coming later or perhaps not at all. For such broad purposes (and most of any enduring educational importance) the original integrated statement vigorously expressed by an able author would seem likely to be more effective than this same matter cut into little pieces and spoon-fed by some programmer.

Since the early days of Binet testing it has been known that, for instance, the children in a fifth-grade class might range in mental age from 8 to 15. Over 20 years ago, Learned and Wood (1938) presented evidence that, on a very extensive test battery aiming to cover the essentials of a college education, some high-school seniors scored above the mean for graduating college seniors, and some of this last group below the high-school mean. A recent study of gifted public school pupils found most fourth- and fifth-graders scoring on subject-matter tests above the seventh grade average, most of the eighth-graders up to the average of the twelfth grade, and twelfth-graders above the average for college seniors on the Graduate Record Examination (California State Department of Education, 1960). A rigid set programming would seem inevitably inappropriate for many if not most in any class; and programming steps easy enough for almost all would be like asking 15-years-olds to keep in step with children of 8.

PROBLEMS OF APPRAISAL, AND OF "PROGRAMMING"

Sundry bits of experimentation by the writer and his students have attempted appraisal of the value of automatization in initial stages of school learnings, always in comparison with the efficacy of the same amount of time spent simply in ordinary study of the same matter. So one college freshman class studied a brief passage from a textbook for 15 minutes while another class spent the same time using a self-instructional device to go through a programming of the passage for self-instruction, then both classes took an objective test covering important points in the passage. Another class studied another passage for 15 minutes while a comparison group studied the passage for 8 minutes, and then for 7 minutes ran through a brief self-instructional test aiming to elucidate major points in the passage; 19 days later, both classes took an essay type examination on the major topics. And in all these little investigations, students

did as well on the end-tests when they had simply studied the matter to be learned, as when aided by the special materials!

This result may obviously be attributed to the inadequacy of these bits of classroom experimentation, or the ineptness of the "programming" of the self-instructional matter. But another explanation is conceivable: the average college student (or able sixth-grade pupil) may have attained sufficient effectiveness in reading expository matter of moderate difficulty that, *initially*, he needs no special help! If he does seem to, perhaps it should be first considered whether the passage was not clearly written, or whether a recall was expected which was so over-detailed as to be unreasonable—as is not uncommon in both school and college quizzes. The writer's investigations have shown automated help indeed valuable for selective and corrective review and sometimes for guiding reading or observing or experimenting. Students so helped did better on final examinations than other students in control sections not so helped (Little, 1934; Pressey, 1950). With such help, able students might even cover a course without any formal teaching, by reading and laboratory work so guided and checked (Jensen, 1949). But this seems essentially different from having detailed programming largely take the place of the reading or other first contact with what is to be learned. At this comparatively early and conceivably sometimes overenthusiastic stage of attempts at automation, might this "null hypothesis" be desirable: that students and teachers can (on the task in question) do at least as well without the special device or programming as with it! Only if results have warranted the rejection of this hypothesis, can development of special devices or programmings, or comparisons of short versus longer steps or objective versus write-in answers, continue without risk of being found fatuous.

In selecting material and preparing it for such guiding and correcting use as just mentioned, with a self-instructional device, to cover a college course, the writer has considered three steps desirable. First, pretests of a short-answer essay-type should be prepared covering the essentials of the course and given at its beginning. Second, questions that are found to present any difficulty should be thrown into multiple-choice form, with the right answer very carefully formulated as elucidative of its rightness, and the wrong answer representing those errors occurring often enough that they seemed to need dealing with. Third, in the next giving of the course, a sampling of the pretest questions should be given both control and automating sections at the beginning and in repeated or paraphrased form in the final examination, which should also include questions not this time used in the self-instructional material. Experiments of this type have shown the sections using the self-instructional device doing better in the final examination, not only on items repeated or paraphrased from the self-instructional test, but even on those not used in it; there was spread or transfer (Pressey, 1950). Further, the most valuable self-instructional items were those "of substantial difficulty, but presenting a clear and significant problem" (Jones, 1954). It was concluded that the revised self-instructional materials should be made up entirely of such items.

All this may seem to the detailed programmer theoretically wrong in three ways: supposedly learning-producing very easy items go out; the elimination of substantial numbers of them leaves so many holes in any initial close-woven tissue of programming that what is left becomes more a series of somewhat disconnected probing questions; and learning seems to result from mistakes—immediately corrected. On this last point, detailed analyses of results from automated and control classes indicated that students who had initially been wrong on an item and been immediately corrected did indeed as a result more often than the controls answer correctly on the final examination several weeks later; in addition, more of the experimentals than of the controls who got an answer right on the pretest stayed with that right answer on the final—the automation apparently confirmed what may have been an uncertain initial right answer and warned against plausible wrong answers which might have been turned to in the final (Jones, 1954; Stephens, 1953). Surely, thus, building self-instructional material of items empirically shown to be instructionally potent seems reasonable. Much has been made of the analogy that automation served like an able tutor. But surely a wise tutor does not all the time lean over a student's shoulder, continuingly asking very easy questions. Rather he sets tasks, asks questions stimulating discriminating thinking, gives the student assurance when he is right, guides him aright when he is wrong, and tries to help him increasingly to proceed on his own. And as to inconsistencies of all this with current learning theory: might it be that learning of meaningful material, by a person already somewhat knowledgeable, in the service of his purposes, goes forward in ways calling for some readjustment of theories having origins mostly different? In the elucidation of very difficult or special matter, or for special groups as young children, detailed small-step programming may be exceedingly useful—but surely not for everybody on everything.

PROBLEMS OF PRACTICALITY

Some years ago, after experimentation evidencing the value, in a course in psychology, of certain self-instructional materials and devices, the writer made all this equipment available to the instructors in that course. They made little use of these aids. And when the writer taught a section, he did not either! Upon consideration, the reason seemed elementary but basic: it was too much trouble! Various materials had to be got out, and coordinated (a wrong test or wrong key for one student might throw a class into a tizzy) and carefully distributed and then carefully collected and recorded and put away, and the total might be two or three hours of fussy work. Significant research may legitimately be expensive in time and money, of course. But if an intent is to come through to something which will in actuality be useful *and* used, then consideration of such practicalities should not be long delayed. A second "null hypothesis" is here ventured as desirable to keep in mind: that the device and program in total takes so much time and trouble (or is so expensive) that in cold fact it is impractical for everyday use by the people and in the circumstances for which it is intended. And more important than the expense is the trouble! Unless some of the current materials and procedures (hundreds of programming charts or frames to be kept in order and used just right) are made a lot less clumsy and easier to use, this second null hypothesis will get them.

Relevant here is the present hassle over the comparative merits of multiple-choice versus write-in or constructed responses to automated matter. The writer is reminded of the somewhat analogous controversy, some 35 years ago, over objective versus essay-type examinations. At that time it was increasingly realized that: (a) in practical fact the life tasks for which the school should educate rather more often called for recognition or choice of the best alternative in a situation than writing a paragraph about it; (b) with ingenuity, objective test forms could be made to cover a greater range of traits or learnings than earlier appreciated; (c) as use of objective forms became habitual and the matter was meaningful, form of response became less important; (d) most students came to like objective forms as saving them trouble (as in writing) and covering more topics more incisively than written responses; and (e) though they might not admit it, most teachers also like objective tests especially because of their convenience. The prediction is ventured that, for largely similar reasons, objective self-instructional matter will increasingly predominate.

Sundry evidence indicates that learning proceeds more efficiently if the learner is kept informed of goal and progress, and it is usually taken for granted that the teacher is aided by having such information about each pupil. If materials for self-instruction are carefully selected, they might be considered statements of goals to be mastered, and a record of efforts to do so a record of progress toward those goals. The writer's students have obtained indications that learning might thereby be increased. Teaching machines which automatically kept a score, and were easily reset so that each student could repeat each self-instructional test until his score was perfect, were liked and appeared to bring about higher grades on objective final examinations than might have been expected otherwise (Little, 1934; Stephens, 1953). When an instructor arranged that students scoring low on initial self-instructional tests were given special individual help by others in the class scoring high, both these extremes were found to do better on the final examination than control students equally low or high but not participating in such a "tutoring" arrangement (Jones, 1954). Since many self-instructional devices can easily be made to yield scores, it seems strange that much current work neglects such possibilities for measurement aiding learning.

The statement is now not infrequently heard that the device is comparatively unimportant; the programming's the thing! Certainly the subject matter for automation must be selected and organized on sound bases. But the full potentialities of machines are now only beginning to be realized. Thirty years ago, the writer demonstrated apparatus that (a) permitted shifting back and forth from feedback to measurement only, thus providing intermittent reinforcement if desired, (b) repeated for each learner only those items causing him trouble, and (c) gave a candy reward for good work. Now Crowder's "tutor" may show in a colored movie how to do a thing, show it again if needed, branch this way or that—in a sense, automatically tailor-make a program to fit each learner. The "Videosonic" devices Keislar is currently using [see Fry *et al.*, 1960, p. 16] employ pictures to develop concepts in young children, talking to them if they can't read. Some devices make it possible for the child to talk back. If only psychologists and engineers and educators (and enough money) can all get together, almost anything is possible. The device must be practical.[2] But to modern engineering, that does not set limits, but only requirements of convenience.

SUMMARY AND PROPHECY

This paper began with certain criticisms of some current work in automation. Machines are not, the contention is, going to make unnecessary either educative contact with reality, competence in reading, sociality, or adjustment to individual differences. They must be kept subject to educational evaluation and curriculum-making. They must be a convenience, not a bother. The teaching machine is no Frankenstein monster, threatening all present texts, materials, equipment, and methods, and even many teachers! Rather, automation should be the greatest help the schools have ever had. Only continuing cooperation between automators and educators is needed to make this so.

---

[2] The device need not be mechanical; it may be a "scrambled book" which not only puts questions and appraises answers but also explains, and directs further work according to need. Alternatively, self-instructional tests (with answer sheets on which a mark in the "right" place is shown to be such by an immediate change of color or other cue) may indicate remedial study references or give other guidance. Such matter may be planned to be self-sufficient, or to be used with conventional instructional materials. As devices become thus increasingly versatile in fostering and guiding learning, and increasingly practical, they should become essential educational tools.

Error Responses and Reinforcement Schedules in Self-Instructional Devices

ABRAM AMSEL

IT IS the purpose of this article to present some reactions to what seem to be some of the premises set forth by Skinner and his colleagues in connection with their use of teaching devices.

Typical of Skinner's current approach to the study of learning, he and his students have shown a great proclivity for building teaching machines and designing programs of instruction; but they have been relatively inactive in the business of providing a set of principles to guide their work and the work of others in this field. There are, however, two articles by Skinner in the published literature which provide something of a rationale for the Skinnerian approach to teaching machines.

In the first published reference (1954) to his interest in teaching machines, Skinner uses the first half of his presentation to discuss, in general terms, the very interesting work he and his students, particularly Ferster (1957), have done with pigeons on schedules of reinforcement. This work is mainly concerned with promoting different rates and degrees of persistence in operant responses by manipulating reinforcement schedules. From this discussion, Skinner moves directly into the area of education (formal schooling) with the implication that the principles derived from the operant pigeon behavior can be employed in setting up and using teaching machines. Skinner seems to feel that educational procedures, and presumably also training courses of most kinds, are deficient in two main respects: (a) They are based upon aversive motivational control, e.g., fear and guilt, and (b) they do not provide for an adequate number of "reinforcement contingencies." There is also the suggestion that somehow reinforcement schedules will be important in the use of teaching machines.

An earlier version of this paper was prepared as an informal staff memorandum in 1957 while Dr. Amsel held a temporary research appointment with the Air Force Personnel and Training Research Center at the Maintenance Laboratory, Lowry Air Force Base, Colorado. He is currently Professor of Psychology at Tulane University.

Skinner's recent article in *Science* on teaching machines (1958c) provides a clear picture of the vigor with which the building of machines and the design of instructional programs have been pursued by him and his colleagues. There is in this article very little space devoted to systematic considerations or presentation of principles to guide the builder of machines or designer of programs. The closest one comes to this is a section headed, "Can Material Be Too Easy?" in which Skinner again emphasizes aversive motivational control [in traditional educational practices] and attempts in his teaching machines to "maximize success and minimize failure."

A good source of information about the rationale of Skinner's approach is a reported interview with Skinner.[1] The passages which comprise the following paragraphs are taken from this report. Italics are added to point up what the present writer feels are important parts of the reasoning.

"Skinner's teaching machines permit the student to work at his own rate and are thus adaptable to wide ranges of intelligence . . . are self-motivating . . . and offer convenient methods for providing reinforcement. Skinner reports . . . students on whom he has so far tried the machines find operating them so much fun that no further reinforcement is needed, and he is not now doing any work on the question of how to use reinforcements in connection with such machines. From his extensive work on reinforcement schedules, however, he feels that the *most advantageous reinforcement schedule for such a purpose* would be a variable-ratio schedule. . . . The typical slot machine pattern, in which there are fairly frequent small rewards and a very infrequent big reward, seems especially well suited to maintain a *high rate of machine operation.*

"Teaching machines, Skinner feels, should teach by a Socratic method of questioning, with a minimum of exposition and a maximum of active participation by the student. *The step between each question or item of information and the next should be very small, so that the student is seldom wrong.* This is important both to maintain motivation and to make sure that the student gets a complete understanding of the material being taught. The material should be presented in a logical progression so far as possible, and in such a way that the student understands what he is learning, rather than learning by rote."

---

[1] This is an interview which Ward Edwards conducted with Dr. Skinner in 1956. It was printed as Laboratory Note ML-LN-56-3 in May 1956 at the Maintenance Laboratory of the Air Force Personnel and Training Research Center, Lowry Air Force Base, Colorado. (See abstract of this paper in Appendix I.)

Two issues emerge which have not had sufficient consideration in the management of education and training. The first has to do with Skinner's feeling about aversive motivational control and maximizing reinforcement contingencies. It relates to Skinner's recommendation that the steps between questions or items should be small, so that the student is seldom wrong. The question is: Is relatively errorless practice always the best way to learn? The second issue has to do with the role of schedules of reinforcement in training situations. The question is: What aspects of educational learning will be affected by intermittent reinforcement schedules, and how will they be affected? These are separate questions involving separate manipulable variables: manipulation of the number or proportion of incorrect responses in a training situation in the first case (this can only be controlled through the difficulty of the material) and manipulation of the number of nonrewards in training in the second (this can be controlled arbitrarily). The remainder of this paper is a discussion of these two classes of variables as they form part of the rationale of self-instructional (educational) devices—or. for that matter, any kind of training or educational system.

ERRORS AND PARTIAL REINFORCEMENT IN TRAINING

It should first be pointed out that varying the number of incorrect responses does not constitute varying the degree of partial reinforcement. There is occasionally confusion on this point, which is not lessened by Skinner's transition from consideration of the operant responding situation in pigeons to the teaching machine situation in humans. The confusion is due, not to species difference, but to the difference in learning situations employed with the two species.

Strictly speaking, the word "incorrect" describing a response implies differential responding and different consequences for alternative responses. The differential response situation—a situation in which a response can be incorrect—has been termed "trial-and-error learning" or "response selective learning." Selective learning implies differential consequences — "correct" (perhaps arbitrarily designated correct) responses are rewarded; "incorrect" or "error" responses are nonrewarded or even punished. At the beginning of such a sequence of correct and incorrect responses occurring in chance arrangement, reward will occur on less than 100 percent of the trials, the frequency of occurrence depending upon the number of alternative responses and the initial probabilities of occurrence of these correct and the incorrect alternatives.

Partial reinforcement implies a single, prepotent response, already dominant in the hierachy—or, at least, other responses are ignored—which is rewarded not consistently and continuously but intermittently. The important factor is that there is inconsistency in reward occurrence but not in response occurrence. Objectively, the same response is sometimes rewarded, sometimes not.

Role of Errors in Acquisition of Skills and Knowledges

Now, let us evaluate the role of incorrect responses in training. As was indicated, Skinner recommends that in the use of self-instructional devices the training be arranged so as to insure the least occurrence of errors, so that the learner will go in very small steps from one correct response to another. He seems to base this recommendation on the following premises: (a) that educational control should not be aversive, (b) that there be a great number of reinforcement contingencies, and (c) that complex behavior is best learned through a series of progressive approximations to the final complex behavior desired.

There are two related and important questions that must be asked at this point by anyone considering seriously the use of teaching machines:

1. Is it clearly always desirable to minimize errors in education and training? Stated otherwise—does elimination of errors in performance always promote the most efficient learning in terms of both the rate of learning and the ultimate level achieved?
2. Does this recommendation of Skinner emerge from some body of experimental evidence—his or anybody else's?

The answer to the second question is, I believe, No. The first question can be answered, ultimately, only through a program of experimentation. It is an important enough question, so that such a program should be undertaken with particular reference to teaching devices of a self-instructional nature. In the absence of a body of experimental data relevant particularly to this question, it is possible to develop a rationale for answering the first question in the negative.

If the answer to the first question is negative, then it should be possible to suggest answers to a modified form of the question: Under what circumstances, or for what purposes, would it be desirable not only to reward correct responses, but also to not reward incorrect responses? Two circumstances suggest themselves, one related to the initial condition of the student or trainee in relation to the training program, the second to the practical consequence of making an error.

FIRST CIRCUMSTANCE. Consider individuals entering an instructional situation which may be represented abstractly as learning to make a class of responses, R_a, to a class of stimuli, S_a. Within those classes a variety of specific associations must be acquired, e.g., $S_{a_1} - R_{a_1}$, $S_{a_2} - R_{a_2}$, and so on. Perhaps this is only preliminary to learning more complex associative chains and interrelations which may be termed skills or knowledge.

Now consider two different initial conditions of the learner taken from a much larger number of possible initial conditions. What differs in these initial conditions is the number of *strong* superthreshold response tendencies which the learner brings into the situation in relation to the class of stimuli, S_a. The two initial conditions to which we refer are schematized in Figure I. Represented here is Condition I, in which the trainee begins with two classes or responses to S_a (R_a and R_b) which are of about equal strength and are very strong; i.e., he starts with strong correct and incorrect associations. The other possible response classes are very weak, so weak that their likelihoods of occurrence are inconsequential. In Condition II none of the possible response classes is strong; all are very weak, and the probability of occurrence of any at the outset is very small.

It can be argued that different strategies for training, with respect to the evocation of errors, are called for in Conditions I and II. In Condition I R_b responses (errors) should be elicited and nonrewarded, because strengthening of the $S_a - R_a$ association alone cannot produce a sufficient differential in strength between $S_a - R_a$ and $S_a - R_b$ so that one can be certain that R_b error will not occur regardless of duration of instruction. In the case of Condition II, it is sufficient to strengthen $S_a - R_a$ associations, because any substantial reinforcement of these associations will make it even less probable than initially that error responses R_b to R_z can occur. In other words, Condition I calls not only for strengthening $S_a - R_a$ but also weakening $S_a - R_b$; in order for $S_a - R_b$ (an error association) to be weakened, R_b (an error) must occur. In the case of Condition II, the training need not involve the occurrence of errors.

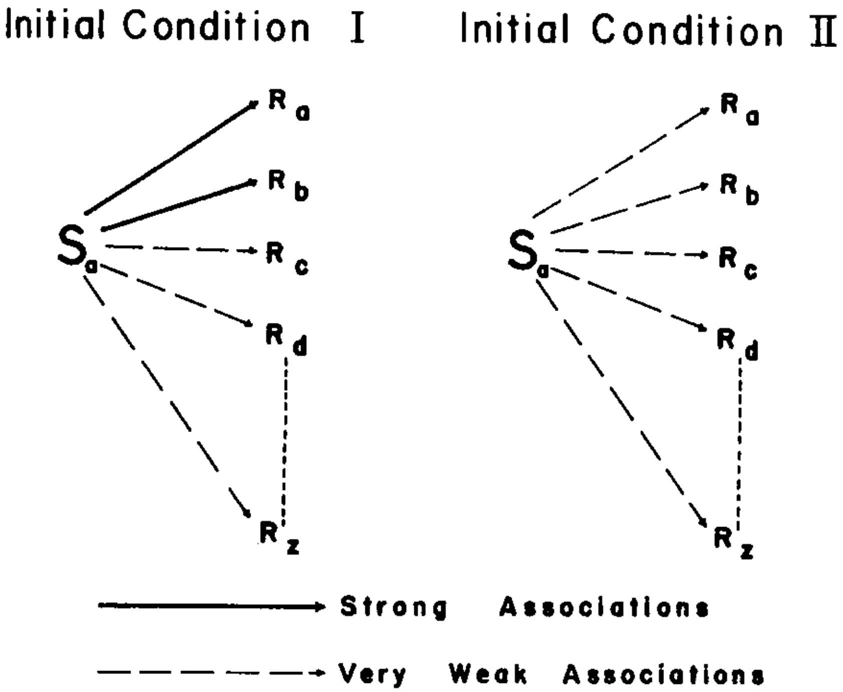


FIG. I. Two initial conditions of the learner with respect to strength of response tendencies associated with S_a. Solid lines represent strong associations; dashed lines represent very weak associations.

SECOND CIRCUMSTANCE. This one involves a different consideration, the consequences of making an error, and is most readily understood, perhaps, in relation to practical training situations. To the extent that consequences are serious, to that extent must training decrease the probability of making errors. Also, to the extent that specific errors are probable, to that extent must training decrease the likelihood of their occurrence. Again, if error response tendencies are strong at the outset of training, it would seem important to elicit them in training and nonreinforce them in order to decrease the probability of their occurrence below some tolerance level. The seriousness of the consequences of making an error should determine which error responses one should attempt to evoke and weaken in training. This has been schematized in Figure II, where incorrect response tendencies (R_i) are arranged in hierarchical arrangement according to strength. The correct tendency (R_c) may have any strength at the outset.

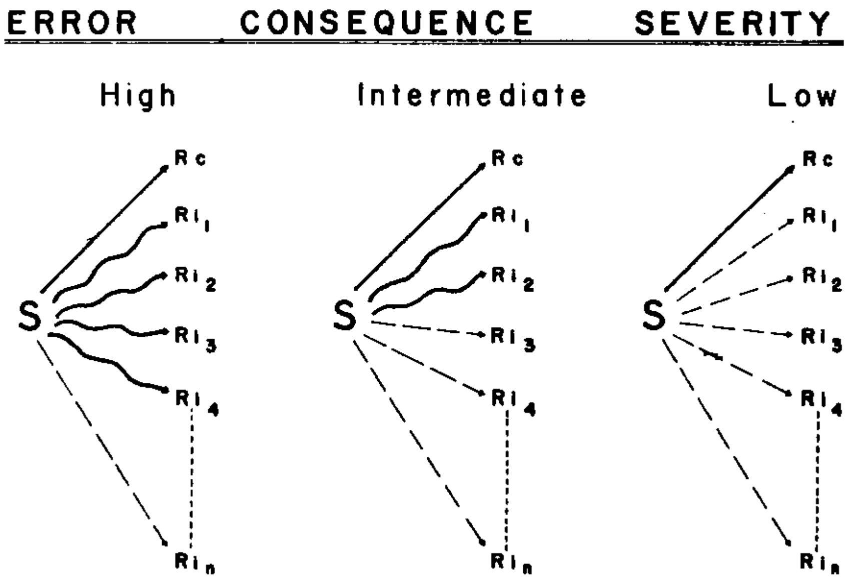


FIG. II. Schematization of the extent to which errors should be evoked and not rewarded in training depending on the consequences of making an error "on the job." Solid lines represent correct (reinforced) associations; wiggly lines represent evoked, but nonreinforced, incorrect associations (errors); and dashed lines represent nonevoked error tendencies.

Role of Reinforcement Schedules in Teaching Procedures

Consider again the teaching-machine situation in which there is a gradual acquisition in small steps of a fairly complex skill or knowledge (pattern of associations). Can partial schedules of reinforcement be employed as a condition of training here? Strictly speaking, they cannot, because the same response is not recurring as in the case of building up rate of responding or persistence of responding of a simple instrumental response.

Skinner and others have been working with two facets of behavior organization, not always clearly differentiated: (a) "shaping" of complex behavior and (b) building up rates of responding (increasing the length of a chain of simple responses) and also building persistence (nonreward tolerance) into the response. The teaching machine situation is an example of shaping of behavior, in that a final complex behavior organization is approached by successive and progressive approximations. This would be exemplified by the following procedure with pigeons: A reinforcement-signalling stimulus is first established, e.g., the sound of opening a door to a food bin, and a pigeon learns to run to the food bin and eat whenever this sound occurs. The experimenter then watches the pigeon and snaps the door open when the to-be-shaped behavior starts to occur (at first by chance). On successive occasions the experimenter waits until more and more of the total response occurs before raising the door; and finally, the pigeon is performing a fairly complex response in a highly stereotyped fashion whenever the experimenter stands by the food-cup door. The response is to run around the perch some given number of times, then to the food bin.

What is important is that in this procedure—in the shaping of a response—continuous reinforcement is utilized. However, *once the response has been shaped,* the persistence of the overall response and increase in its rate of occurrence can be promoted by instituting a schedule of reinforcement *for the overall response,* i.e., by not reinforcing it every time it occurs but according to some ratio schedule.

If an analogy can be drawn between response shaping or formation in pigeons and the acquisitions of a relatively complex response (skill, knowledge) in humans, then implications for the use of teaching machines and other self-instructional devices are clear. The gradual acquisition of a behavior pattern, skill, or knowledge by increasing in small steps the complexity of the response should not involve any schedule of partial reinforcement. Partial reinforcement schedules may have a place in training of skills and knowledges, but this place is probably not in the shaping of a relatively complex behavior or associative pattern, because shaping involves a gradually changing and emerging response pattern, not a single relatively unchanging response.

Schedules of reinforcement (patterns of intermittent reinforcement and nonreinforcement) may, however, be important in training which employs automatic devices in two respects: (a) to maintain an optimal level of device operation in the acquisition of the shaped behavior (skill or knowledge), and (b) to increase the persistence of the behavior once it has been shaped. But something more should be said about each of these proposed uses of intermittent reinforcement:

1. Device operation should be differentiated from acquisition of the shaped behavior, implicit or explicit, which is the purpose of a program of training. Certainly, the device must be operated in order for learning to occur; however, conditions which might increase persistence of device operation might have no effect, or even a negative effect, upon the learning which the device is to promote. For example, the slot machine analogy—"fairly frequent small rewards and a very infrequent big reward"—described by Edwards might produce a very high rate of machine operation but not a very high level of correct differential reaction to the materials presented by the device. The self-instructional learning situation is partly, at least, an operant-rate-conditioning situation; and although there is no wish here to conclude that the operant aspect of the situation must necessarily interfere with the learning of material exposed by the machine, neither should it be concluded that achieving a high rate of persistence of machine operation through reinforcement scheduling will necessarily result in either a high level of knowledge acquisition, simply because the device has exposed material many times, or high motivation for the acquisition of the associations involved.

2. The second possible use of reinforcement schedules in training is related to the earlier discussion of the distinction between shaping of behavior, on the one hand, and building up rate and persistence of the shaped behavior, on the other. In the case of training, this boils down to a distinction between the acquisition of a knowledge or skill and the acquisition of persistence in the application of that knowledge or skill. In most types of training the use of schedules of reinforcement should have consideration only after the behavior, which the training was designed to promote, has already been shaped. Then, it becomes a question of whether persistence in the face of nonreward (negative indications) should be built into the shaped behavior. If persistence of this kind is desired, it can only be acquired through experience of failure and learning to continue responding despite failure.

For example, training of a skillful response, such as might be involved in making a fine adjustment, should probably involve partial reinforcement schedules to establish persistence, a behavior attribute which would insure the eventual making of the optimal adjustment. On the other hand, partial reinforcement would not be recommended in training to trouble shoot, primarily a "problem-solving" activity. In fact, every attempt should probably be made to build consistency into such a training program. In training of problem-solving kinds of behavior, partial reinforcement would result in preservation—persistence of behavior, which does not solve the problem. One way of stating the principle involved in this distinction might be as follows: The use of partial reinforcement schedules is indicated in training where negative feedback stimulation (indications of the inadequacy of the behavior) should result in continued refinement of the same kind of activity until its adequacy is signalled. Continuous reinforcements in training is indicated where behavior versatility is desired and where feedback stimulation indicating performance inadequacy should trigger an immediate change to a different kind of behavior.

SUMMARY

The strong, current interest in procedures attending the use of self-instructional devices would seem to call for systematic analysis of some of these procedures by theoretically oriented learning psychologists. The present paper is a preliminary examination of (a) the Skinnerian rationale for relatively errorless learning and (b) the applicability of reinforcement scheduling to self-instructional training, or any other.

A rather simple analysis provides an argument that it is not always desirable to minimize errors in education and training. It is argued also that the use of partial or continuous reinforcement schedules should be decided by the nature of the activity the training is to promote, partial schedules being mainly suitable for building persistence into a relatively unchanging behavior pattern.

Some Issues Concerning Devices and Programs For Automated Learning

A. A. LUMSDAINE

INTRODUCTION

THIS PAPER consists of two general parts. The first part deals broadly with some major issues and classes of variables in the development of automated or "programmed" methods for individual instruction. This will provide perspective in which we may view some illustrative problems that are discussed more explicitly in the latter part of the paper. This introductory portion may also help provide a frame of reference to which some of the other papers of this volume can be related.

PART I—GENERAL CONSIDERATIONS

Some Common Features in Automated Learning Devices

All the devices we now call "teaching machines" utilize some variation on what can, in a loose sense, be called a "Socratic," or tutorial, method of teaching. This method incorporates three important features, identified in the left-hand column of Table 1.

As is indicated at the right in Table 1, these three features, in effect, implement: first, the dictum of Guthrie and others that the student will *learn* only what he *does;* second, an emphasis on feedback, reinforcement, or knowledge of results; and third, the desirability of making suitable provision for individual differences in learning ability.

An abbreviated form of this paper was presented at a symposium at the September 1959 convention of the American Psychological Association under the title "Response Cueing and 'Size-of-Step' in Automated Learning Programs." Parts of the paper were prepared in connection with research at the American Institute for Research sponsored by the U. S. Office of Education, the U. S. Air Force Cambridge Research Center, and the U. S. Air Force Office of Scientific Research.

Table 1. COMMON FEATURES OF TEACHING MACHINES

FEATURE	PRINCIPLE OR DICTUM REFLECTED
1. Continuous active student response to questions or exercises	1. The learner learns only what he does.
2. Prompt feedback and basis for correction of errors	2. Learning is benefited by prompt "reinforcement" and/or "knowledge of results."
3. Individual students proceed at their own rate	3. Individuals differ greatly in rate at which they can learn and perform.

Most psychologists will see in these three features a potential basis for providing highly efficient conditions of learning, and it is not surprising that these features are the ones shared in common by virtually all teaching machines thus far proposed. Recognizing these communalities among different approaches to automated teaching is a useful background for examining some of the important *differences* among various approaches. What then are some of these major differences, and what are some of the issues important in the study of learning to which they point?

Some Major Issues

For a first categorization at the level of broad areas, we may identify three, or perhaps four, major issues. These are summarized in Table 2. Let us look at them briefly.

Table 2. SOME GENERAL ISSUES CONCERNING AUTOMATED LEARNING METHODS

ISSUE	ILLUSTRATIONS
Issue No. 1: What is the role of automated teaching devices in an instructional plan?	1. Type of instructional goal 2. Relation of teaching-machine programs to other forms of instruction
Issue No. 2: What are the major factors that may be mediated primarily by features of the presentation-and-response **device** (teaching machine)?	1. Form of student response 2. Display capabilities 3. Type of feedback from device to student (and vice versa)
Issue No. 3: What are the major factors associated primarily with the instructional material (**program**)?	1. Use of "small steps" 2. Form of "cueing" or prompts used 3. Logical organization of program

The first of the issues identified in Table 2 is such a basic one that it is all too easy to overlook when one becomes engrossed in the fascinating details of automatic devices and teaching machine programs. Namely, what role is to be played by automated teaching devices for presenting so-called programmed instruction within the framework of a total plan or scheme of instruction? This broad issue includes not only a question of the categories of behavior desired as instructional outcomes, but also the relation of automated teaching programs to other forms of instruction.

The second issue to which many students and proponents of automated learning methods are addressing themselves is: What should be the characteristics of a teaching machine—in terms, for example, of its response apparatus, its display capabilities, the forms of feedback or correction it can provide to the student, and the reciprocal effect of right or wrong responses by the student on what is registered or presented by the device?

However, it is on the *third* issue identified in Table 2 that many of the by now numerous and enthusiastic workers in the automated teaching field are increasingly concentrating their main attention. This issue is conceived to be less a question of the device itself than of the kind of program that the instructor feeds into a device. At least some of the major questions that arise here are more or less independent of the nature of the device. They concern various aspects of the size of the steps by which a student progresses through a subject-matter program, the form of prompting or cueing with which he is provided, the information he is given to correct or reinforce his responses, and the logic of the way in which a subject matter is organized and sequenced in developing a program.

So crucial does the importance of programming appear to be for the successful exploitation of a teaching machine approach that a number of investigators have lately taken a rather extreme position concerning it. "Programming," they say in effect, "is everything: the characteristics of the device are of small importance, if any." About the *converse* of this statement there can be little argument, at least at the point of practical application. Thus, the best teaching machine in the world is rather useless unless we have adequate programs to put into it (and so far as full courses go we have, indeed, only a very few at the present time).

Table 3. SOME SPECIFIC SUBISSUES RELATED TO THE ROLE AND DESIGN OF TEACHING DEVICES

MAJOR ISSUES	SPECIFIC SUBISSUES AND/OR VARIABLES	EXAMPLE
Table 3-A Role of automated teaching devices (Issue No. 1)	1. ("Ends") What are the desired outcomes of instruction to which the automated teaching device is to contribute?	1. Identification learning (paired-associates); procedural or sequential learning; "substance" learning—principles, etc.; problem solving and creative thinking
	2. ("Means") What part is the automated teaching device to play (vis-a-vis other elements of the instruction)?	2. Sole vehicle of instruction; major complementary component; minor adjunct; etc.
Table 3-B Device-linked factors (Issue No. 2)	1. Form of student's response	1. Construction vs. recognition; "free" vs. constrained or discrete; large vs. small number of possible choices; etc.
	2. Basis for feedback/correction to the student	2. Automatic vs. judgmental scoring; correct answer necessarily revealed vs. not; correction at end of lesson, "frame," or constructed-response element (e.g., each letter or digit); etc.
	3. Effect of student's response on the program	3. Item drop-out; program branching, item by item or section by section; program variation based on single response or on performance over a series; program variation in content, speed, etc.; automatic vs. student-mediated branching
	4. Visual display characteristics	4. Controlled sequencing of large display (including pictorial) vs. spatially constrained; fixed vs. changing display within frames; special provisions for prompting cues; etc.

But the position that we should concentrate our attention almost *solely* on developing programs and rationales for programming seems less obvious. It appears to the writer that we need to be quite concerned also with the properties of existing and future *devices* that may be used as the vehicle for presenting these programs to the student. These device features may determine or limit the properties of the programs, and the ways a device can interact with the student's responses. It thus seems ill-advised to disregard the character of the machines in our proper concern for a primary emphasis on programs.

Outcomes

Role of Automated Methods in an Instructional System

Let us return briefly to the two main aspects of the first issue. These are outlined in Table 3-A. They are, first, the type of learning outcomes sought—a question of the desired "ends" to be sought—and, second, the role of the device in a total instructional plan in relation to other elements of teaching—a question of the means to be employed.

The factors identified in Table 3-A may be thought of as questions of educational curriculum and administration of instruction. Obviously, however, they also have a bearing on the research design appropriate in investigating learning factors as these are implemented at the applied level in practical instruction. Some of the specific questions which are subsumed here may be asked either as general problems of educational strategy or as tactical decisions in the use of a particular device or program for a specific subject—say, for example, a high-school algebra program presented by a modified Skinner machine.

TYPES OF OUTCOMES SOUGHT. In using a teaching machine, we may be concerned primarily with the pervasive but seemingly simpler forms of specific verbal learning that are easily describable in terms of paired associates. Or we may, for example, have to deal with acquisition of usable concepts and principles that are ordinarily tied to some form of connected discourse, where substance learning and interchangeability of symbolic responses are deemed essential. Again, we may concern ourserves with the learning of essentially sequential patterns, either verbal or nonverbal, as in memorizing a prose passage, poem, procedural list, or some series of motor acts to which we want a symbolic or perceptual sequence to transfer as a mediating instrumentality. Finally, we may be interested in teaching directly some criterion performance. This may be either a specific one involved in a motor act (like key punching or typewriting), or it may be a more generalized one which we would describe as the use of concepts and principles in such intellectual skills as malfunction diagnosis or other forms of problem solving. Such distinctions as these among kinds of learning outcomes which may be the object of endeavor in programming a teaching machine need to be considered in discussing research issues, because the kinds of tasks to be mastered can have an important bearing on the kinds of devices and programs that may be optimal for various purposes.

Many psychologists as well as most educators have an understandable preoccupation with verbal behavior when they are concerned with human learning. But in addition to some important differentiations that need to be made among classes of verbally mediated responses, the students of nonverbal learning need to be given their due attention also. As for the background that may be involved from studies of human learning in the psychological laboratory, perhaps the pursuit rotor runs a close second to the nonsense syllable presented on the drum in the literature of experimental psychology—and, as a matter of fact, both devices meet the basic teaching machine criteria of continuous active response, feedback, and individual pacing—even though the responses they teach may be largely useless and far removed from the kinds of material that most teaching machine programmers are currently concerned with.

Role of the machine. The second of the two distinctions introduced in Table 3-A concerns the role of a teaching machine program in relation to other instruction. One bold and by no means necessarily impractical role is to place sole or primary dependence on a question-and-answer program presented by some form of machine or device to carry the main burden of instruction for a course or at least for certain subunits of a course.

A second kind of role is one in which the teaching machine program forms a follow-up, complement, or adjunct to other instruction. The original multiple-choice programs of Pressey and his associates were of this type. It was not presumed that the teaching machine program would be used until after considerable prior study or lecture instruction, and the quiz-like items and their sequencing were conceived as having follow-up review and testing functions. More recently, experimentation has been initiated with programs that are planned as a complement to basic lecture-demonstration teaching presented by film or television. Here their role can be that of presenting and correcting an extensive program of exercises to be done by students in a study period or as assigned homework. Finally, considerable attention has been given, capitalizing on what we may term the pinball philosophy of instruction, to making machines available for voluntary use by students in their free time. The optimal characteristics of devices and programs (and related research issues) may obviously differ, depending on which of such roles a teaching machine program is expected to play.

Device Characteristics and Their Relation to Programming

One thing that emerges from the foregoing distinctions is that such factors as form and diversity of student response which a teaching machine must permit, or the kind of response language it must understand and be able to react to, may in some cases have to differ greatly with the function the device is to serve. With this notion in mind, let us examine the response repertoires provided by some existing teaching machines.

Four major characteristics with respect to which some of these devices differ are listed in Table 3-B. These are: (a) form of student response, (b) basis for feedback/correction to the student, (c) reciprocal feedback to the device and its program based on the student's response, and (d) visual display characteristics.

Now let us look at Figure I to illustrate how these and other related differences are exemplified—in a number of devices. The devices at the top of Figure I all employ freely constructed responses made by the student. The devices at the top left (Skinner, Porter, and others) are designed primarily for "continuous discourse" verbal subject matter. Correction or confirmation of a written response is made possible when the student causes the device to display the correct answer.

Judgmental scoring on the basis of comparison with a model is also employed in the kind of device illustrated at the top right, although the form of response and kind of learning required are quite different (Lumsdaine, 1959a,d). Here, a motion picture or slide presentation with accompanying recorded oral directions is used to teach sequential procedures. As in the verbal learning machines just above, the learner can proceed by very small steps, here consisting of short demonstration segments, which he can readily imitate and then check the step just taken against the pictorial model on the screen.

The external manipulanda employed are, of course, much

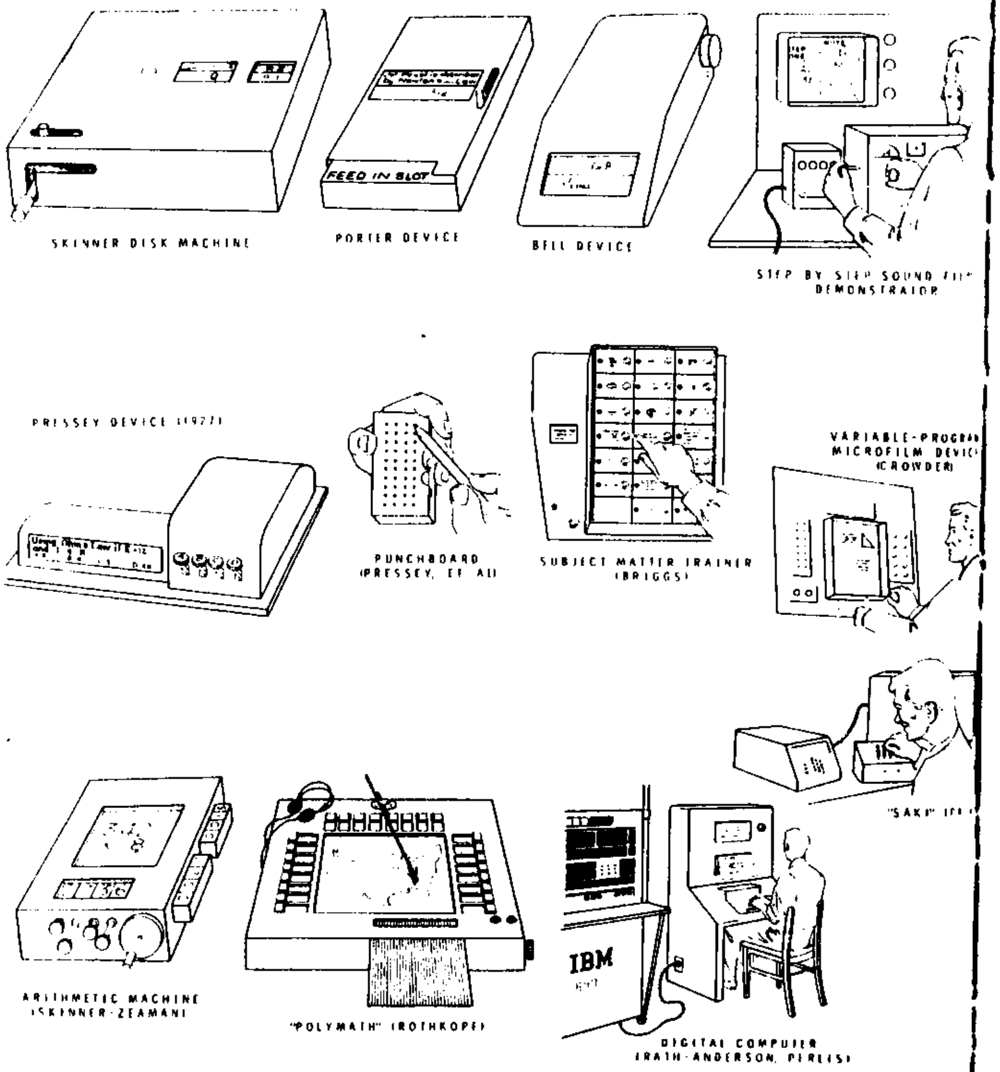


FIG. I. Teaching machines illustrating various characteristics

more complex than with the verbal machines, where only a pencil or pen is required. However, the essential pattern of feedback is basically similar, inasmuch as it consists of comparing *some* form of freely constructed response judgmentally against a visual criterion. Also, the mediating symbolic responses may involve similar implicit verbal processes even though the performance and learning outcomes are quite dissimilar.

In contrast to the devices at the top, the devices near the center of Figure I (Pressey, Briggs, and others) differ in several respects in the response form they employ and the type of feedback they provide to the student. First, the apparatus is self-contained (not requiring even such minimal external apparatus as a pencil or a pen to make a response). The student simply pushes a button to indicate his answer.

Second, the response repertoire is restricted to discrete choices rather than allowing free responses. This has at least two important implications. In the first place, the device can be said to involve *selective* rather than constructive responding. However, this distinction can easily become blurred if the number of choices is large (consisting of up to 20 choices as in the Briggs Subject Matter Trainer, or even more in other devices). The distinction is also blurred where the criterion response desired *is* actually that of selecting or identifying, or where a "freely constructed" response can be coded or analyzed into a finite number of possibilities.

In the second place, the use of discrete responses permits automatic scoring or right-wrong discrimination rather than dependence on judgmental comparison as the basis for scoring, correction, or confirmation. Because of this, correction-feedback or scoring may be accomplished without directly revealing the correct answer. Thus the student's selection of a wrong response may (depending on how the device is constructed and programmed) score an error and require him to make one or more further attempts before he can proceed to the next question in the sequence.

Thus there is something of a trade-off in flexibility between these two classes of device. The devices at the top of Figure I permit more variety in readily-scored response form (and may admit of response equivalence in terms of synonyms or other alternate responses). But, on the other hand, more flexibility in manner or confirming or correcting a particular response is afforded by the discrete-choice devices, since they do not have to reveal the correct answer in discriminating an incorrect response.

It can be further noted that use of discrete choices for scoring and feedback need not involve the alleged limitation of purely recognition responding: The student may *first* "construct" his response (in writing or mentally) and *then* "code" this "free" response in terms of category-choices which can, if desired, be withheld from him until he has completed his initial free response. A simple modification of existing multiple-choice devices would provide for this technique, which has also been employed in some unpublished Air Force studies in using motion pictures or slides as testing devices.

At least one important feature is shared in common by the devices in the top two rows of Figure I—or at least by all but the simplest of the latter (punchboard, which is used with an accompanying sheet to present a series of questions). This is the fact that they provide stimulus control of the question elements in a program, so that the student is prevented from looking backward and forward in the series while responding to a particular item. In these particular devices it may also be noted that the length of the verbal stimulus displayed for each "question-frame" in a program sequence is spatially constrained, permitting only a short verbal question-statement. More extensive stimulus material (e.g., any sizeable pictures or diagrams) must therefore be presented through adjunctive display panels, for which the same degree of stimulus control is not necessarily available as for the basic question-frame. This disadvantage is avoided by using projected microfilm, as in Crowder's "Tutor" (p. 297) and other recent devices, e.g., Wyckoff's device (Teaching Machines, Inc.).

The devices depicted at the bottom of Figure I combine the features of "constructed" response with discrete and automatic wrong-answer discrimination (and the attendant characteristics of this feature as just identified). The simplest of these is the special-purpose arithmetic machine (Skinner-Zeaman) at the lower left, in which sliders and dials are used to "set in" an answer to a numerical problem. More complex devices, employing either keyboards or sliders for setting in the answer and permitting an extended alphabetic as well as numeric repertoire of responses, have also been constructed.

An experimental device of this sort (Rothkopf's Polymath) has an extended response repertoire which permits extended multiple-choice responding and also an alpha-numeric, constructive-answer input through sliders similar to those in the preceding device. In addition, through use of criterion patterns employing printed circuitry on the back of the display sheet (not visible to the student), two-dimensional positioning or tracing responses can be automatically entered into the discrimination-scoring apparatus and memory of the device (for example, such responses as tracing a route on a map or a signal through an electric circuit diagram displayed in a particular "frame" of the program). The utility of such extended response-registering apparatus for various kinds of instructional requirements represents an interesting field for investigation.

The most complex of the devices that can automatically score a full alphabetic and numeric repertoire consists of electronic digital computers coupled to suitable keyboard input and print-out devices. The complex degree of instrumentation represented by computerized devices permits the incorporation of two additional features with respect to flexible correction-confirmation and program sequencing. The first feature is correction-confirmation provided for each serial element of a complex verbal or numeric response, letter by letter or digit by digit. However, this can also be accomplished more simply by microfilm devices using special forms of response coding, such as the photoelectric response-matching machine recently devised by Wyckoff.

BRANCHING PROGRAMS. A second feature of such devices is that the great versatility of the computer frees the teaching machine from the constraints of simple predetermined sequences. It thus permits the instructional use of contingent branching sequences—in much the same way that a skilled human tutor can adapt his program of questioning and discourse differentially to the response of the individual student and the needs which these responses indicate.

Thus, for example, on the basis of a particular error or pattern of errors, the program presented to the student may be caused to revert to a remedial review or prerequisite subprogram, or, on the basis of near-perfect performance which could result in unnecessary repetition of already mastered material and lead to boredom, the student may be jumped ahead to any desired, more advanced portion of the subject matter or to supplementary alternative material (as directed by program-presented branching instructions to the student or as automatically determined by the control characteristics of the program and device, with no human intervention).

Fairly complex branching sequences may be more simply mediated by semi-automatic means—for example, in the kinds of programs described by Crowder (1959) for use in the variable-program microfilm projector shown at the extreme right of the center row in Figure I. This may be done through what Crowder calls "intrinsic programming." In this arrangement the student's answer to the question can be punched into the machine to indicate his selection of alternatives in answering a conventional multiple-choice question. (Alternatively, he could punch in an answer in accordance with a numerical code assigned to a much larger range of possible choices, e.g., checkpoints in a circuit or locations on a map.) Thus the prearranged program consists of a branching tree covering a large number of anticipated response sequences that the programmer thinks may be called for to adapt to individual differences among students. A quite similar form of branching sequence can be effected much more economically through the use of Crowder's scrambled-book arrangement, or "TutorText" (p. 286 ff.).

The utility of both simpler and more elaborate forms of branching as compared with other forms of programming is an interesting and complex question on which research is needed. Superficially, their potential advantages would seem to be most obvious in presenting problem-solving exercises which inherently comprise contingent branching sequences. The value of such highly flexible item-by-item branching seems less clear where concepts need to be built up by easy steps or where simple identifications or other paired-associate materials are being taught. In fact, if the steps in a program are made sufficiently easy, as advocated by Skinner and others, providing for branching alternatives would seem a waste, since in such a sequence the programmer tries to keep the student from ever making more than a very few errors. If he is highly successful in this attempt, the frequent alternative pathways provided as consequents for errors might be too seldom used to justify them.

Two other forms of program variation not involving item-by-item branching capability should be identified here. The first of these is the drop-out of "mastered" or already correctly responded to items in a moderately long sequence, requiring repetition of only the incorrectly answered items. Determining the role and importance of this feature, which is provided, for example, in Skinner's disk machine, is also a problem for some needed research. (See Rothkopf, page 324.)

Another practical form of branching involves choice of alternative subsets, of a small number of items each, at the end of any one subset—depending on the student's performance on that set. This is readily achieved in any of several devices which contain relatively short sets of items in physically separate "small" packages. For example, in Porter's device (1958c) sets of five items on a sheet can be fed into the device one at a time, and the choice of each succeeding set can, if desired, be elected in accordance with instructions that take into account the student's performance on the sheet just completed.

The numerous problems in connection with program branching or sequence variation, which the differences in device capabilities above identified point up, are of great interest as foci for research. This applies both to device characteristics and to programming approaches. It seems clear that *some* form of branching may be very desirable, if not entirely essential, for a variety of applications. The study of *what form* of branching is to be attempted in programming for various kinds of instructional outcomes calls for experimental comparison of alternatives in which economic as well as psychological considerations need to be considered.

A much less expensive alternative that is nearly as satisfactory in terms of effectiveness of learning may be preferable to a method requiring highly expensive instrumentation or extremely elaborate contingent program sequences. The requirements for branching sequences within a program may also be greatly reduced if it is assumed that a teacher will be present who can take care of certain contingencies not built into the program.

Devices for Training of Sequential Performances

With one exception (the procedural trainer at the top right) the devices in Figure I discussed thus far have been conceived primarily for the teaching of verbal or conceptual information. Here the sequence of instruction is based primarily on pedagogical considerations rather than on inherently sequential requirements for the behavior which represents the objective of instruction. Objectives of training which include the learning of inherently sequenced performances may impose special requirements.

Some properties of one additional kind of device should be noted in relation to the teaching of this kind of performance. This is the teaching machine shown at the bottom of Figure I, which was designed to teach card punching. This device (called "SAKI" by its inventor) has many interesting properties described by Pask (1959b). Here we wish to note only two simple but noteworthy characteristics which are of general interest in relation to other devices.

The first characteristic involves the fact that the electronic feedback mechanism provides for branching in the content of the practice materials that are shown to be in need of differing amounts of practice. This is done not just on the basis of a single response by the student, but rather by integrating his performance in terms of accuracy and speed over a period of time. The second notable characteristic is that the device "paces" the student and is capable of adapting to and of influencing what amounts, in effect, to his "level of aspiration." It does this by adjusting the speed at which he must correctly punch the keys in order to satisfy his mechanical tutor (and obtain whatever attendant reinforcement is thereby provided).

PART II—SOME SPECIFIC QUESTIONS FOR RESEARCH

From the foregoing analysis of device characteristics, it seems quite clear that, even within any defined set of instructional objectives, a host of specific research questions is generated by such differences as those in the response repertoires, display characteristics, and branching capabilities of various teaching machines—and also by the differences in programming philosophies to which these differences are partially linked. Several major classes of questions that arise in the programming of material are outlined in Table 4. The remainder of this paper will discuss the first two of these questions in somewhat more detail.

Table 4. SOME MAJOR QUESTIONS CONCERNING PROGRAMMING OF MATERIAL

MAJOR QUESTIONS	EXAMPLES
1. "Size of step"	1. Length of "frame" (number of words or sentences); difficulty of giving correct answer at that point in program; complexity or length of response; number of responses to questions before correction/feedback or reinforcement
2. Content of prompts or manner of cueing	2. Contextual (semantic or formal) within frame; cues from immediately preceding frames; cues from preceding "trial" or cycle; physical variation in clarity of special prompts
3. Content of correction or answer-frame	3. Right-wrong only; reveals correct answer; gives additional prompts if incorrect; explains why wrong or gives other explanation; etc.
4. Logic of program sequencing	4. Uniform vs. contingent sequences, interacting with type of material and with step-difficulty sought; principles to applications, or vice-versa, or mixed; spacing and form of review; method of integrating part-learning; etc.

"Size of Step" as a Factor in Automated Teaching Programs

The first of these is the question of the "size of step" used in a program. We have already seen some of the ways in which this may interact with the characteristics of the presentation devices.

The important principle that mastery of a complex subject should be built up by fairly small steps is a deceptively simple one, even though its essential soundness for many purposes may be taken more or less for granted. For one thing, the meanings of the terms "step" and "step size" have an awkward way of shifting, depending on the context in which they are used. Some of these differences in meaning are suggested in the top block of Table 4.

The notion of "size of step" may refer to the difficulty (and hence probable correctness) of giving a correct response to any item. Also, it may refer to the difference or *increment* in difficulty between the concepts, relationships, or terminology between successive steps or sets of steps in a program. For example, it is a large and logically difficult step from the elementary information that electric current flowing through a fine wire heats it up, in the first two items of the 35-item sequence reproduced by Skinner (1958c), to the relationships between incandescence and emission arrived at by the end of this sequence. But it is a very small step, which the student takes easily, from any one item in the sequence to the immediately succeeding item.

It may be noted that there is also a dual meaning within this primary sense of step size—namely, between the dependent variable of *difficulty* (objectively indicated by proportion of correct responses), for example, as distinguished from the set of independent variables on which this resulting difficulty depends. The latter can, of course, be manipulated by the kinds of prompting and other programming techniques employed in bridging the gap between successive frames of the program.

A second sense to which the notion of "size of step" may be applied is an entirely different one; namely, what is the size of the *response segment* which the student is required to perform in response to the question contained in each frame of the program? For example, in a verbal-learning sequence this may be a single letter, a word, or several words. More particularly, we may ask about the magnitude of the required response in terms of the number of elementary components that are emitted by the student before a response is corrected or confirmed. Depending on both the device and the programming practices, this may occur one letter or digit at a time (as in the Rath-Anderson IBM-650 binary arithmetic program referred to earlier or in a device recently developed by Wyckoff).

More commonly, a longer segment is employed before differential reinforcement is provided. This may comprise a "package" answer—such as identifying the correct category for a multiple-choice question in a Pressey-type device. Alternatively, it may be the writing in of a word or two in a Skinner free-response machine or setting in a several-digit answer to an arithmetic problem. In the extreme case, as in certain uses of the branching sequences devised by Crowder, the student may be allowed to proceed for a number of frames before he is corrected, so that the "step" he must retrace is actually a sequence of possibly quite lengthy steps.

A third sense in which "size of step" sometimes gets used is perhaps better described as "size of frame." One aspect of this is the interval between successive questions to which the learner is asked to respond. This is most commonly only a phrase or a sentence or two (necessarily not much more, as it is limited by the physical constraints of the aperture) in the Skinner-Porter-Pressey types of device. But in other cases (such as some of Glaser and Homme's programmed text frames) it may be a *number* of sentences, or a paragraph. In the scrambled-book or microfilm projection sequences devised in accordance with Crowder's approach, it may even be several paragraphs.

In effect, here we have deliberately incomplete programming in which several steps are combined into a single step. Dependence is thus placed on implicit rather than overt responding until near the end of the frame. The extent to which this is done is, of course, correlated with the physical frame-presentation capacity of a device and the consequent spatial limits on extent of verbal and pictorial material over which it can achieve controlled sequencing.

In addition, there are important variants in the kinds of difficulty-gap referred to by step size that are determined by the *nature of the learning outcomes* sought. Progression "by small steps" in learning a sequential procedure or a batch of serial verbal material offers a possibly instructive analogy with the "small steps" used in building up conceptual knowledge in the verbal materials with which Skinner's followers have largely concerned themselves. For example, extremely short and easily followed segments of a serial performance may be presented and performed one at a time (i.e., part-learning) or much longer segments (i.e., whole-learning) may be used. And the longer steps may be used either from the outset, or there can be a transition to increasingly lengthened steps as learning progresses through repeated practice sessions (or "trials").

As a matter of fact, one of the few pieces of solid experimental evidence treating variations in length of step analytically comes from the experiments of Maccoby, Sheffield, and Margolius with step-by-step procedural demonstrations. These experimenters found that short steps were better than long ones, at least initially (see Maccoby and Sheffield, 1958). However, they also concluded that self-regulation of step length (by the more able students, at least) leads to their choosing increasing step sizes as learning progressed, with resulting performance superior to that obtained with constant small step sizes. They further concluded that an optimum pattern for average students would be enforced progression from shorter to longer steps. It may be noted that, in this case at least, the size of step (or length of presentation segment before practice) has an almost perfect inverse relationship to the immediacy of the cueing or prompting provided to the student.

Another study (see Evans, Glaser, and Homme, 1959) varied the number of steps required to get through programmed lessons on number theory and music fundamentals. Smaller steps required less time per step (but more total time), and fewer errors occurred in the course of the learning. Beyond a certain point, however, increasing the number of steps in a sequence did not result in improved performance. Coulson and Silberman (1959b), using a Harvard program on behavior analysis, varied step size by deleting certain steps from the program to yield fewer steps but with larger gaps separating them. Smaller and more numerous steps gave better learning scores, but this result is difficult to interpret because (a) the small-step program took more time and (b) it contained more material.

It seems apparent that resolution of the question of "size of step" actually involves a complex of factors which need to be analyzed as a basis for clearer theoretical formulation of their separate functions, as well as their probable interactions in practical learning programs. We may conclude that not only more experimentation, but also better formulation of factors is needed to permit us to speak more precisely about the advantages of "proceeding by small steps" in automated learning programs.

Response Cueing and the Use of "Prompts"

One of the crucial factors in controlling error tendencies, and thus keeping successive steps between frames in a program small (in the sense of easy progression and low frequency of wrong responses), is the use of prompts or cueing devices. Such prompts and some of the factors they involve are familiar to most students of teaching-machine instruction from the discussions in published articles by Skinner (1958c, 1959b) and others. Prompting techniques are also—on an implicit basis, at least—the everyday stock-in-trade of nearly anyone who is successfully engaged in the practical training of either human or infrahuman species.

Essentially what is done is that prompts are chosen which are initially adequate cues for a desired response, and these adjunctive unconditional stimuli are then tapered off, withdrawn, or (to use Skinner's term) "vanished," as learning progresses. In most programs of instruction constructed to date for use in automated teaching devices, the prompts have been provided primarily by clever and laborious salting in of words and phrases that provide formal or thematic prompts in the current frame of the program or in preceding frames. Some of the ways this is done are illustrated in Table 5. In devices that provide for recycling of program segments, *the memory traces* from a preceding cycle or trial (a set of perhaps five to 30 items) might also be considered to furnish one of the conscious or unconscious cues used by the learner.

Table 5. SOME CONTEXTUAL OR SEMANTIC CUEING TECHNIQUES USED WITHIN FRAMES OF TEACHING MACHINE PROGRAMS

---

1. Partial presentation of a word, with omitted letters to be filled in—letters present varying from nearly all present (maximal cueing) initially to all, or nearly all, absent (minimal cueing) terminally
2. Similarity of ideas, calling for a response that is provided in a similar context—e.g.,
   Just as smoke rises, warm air will also ....
3. Similarity of grammatical construction—e.g.,
   The higher the temperature, the faster the molecules move; the lower the temperature, the .... they move.
4. Constriction of the range of response by grammatical construction, e.g.,
   The throttle is advanced just .... the ignition is turned on.
   (requiring a temporal word such as before or after)
5. Similarity of word roots with similar meaning used in a preceding "frame" or earlier part of the same frame—e.g.,
   A candle flame is hot; it is a (n) .... source of light.
   (Desired response: "incandescent")
6. Obvious transpositions
   Gross profit less overhead equals net profit; so to get net profit you subtract .... from ....

An additional form of cueing of considerable interest consists of giving the learner a partial or faint representation of the response cue, which directly, but incompletely, suggests the desired response. Variation in effective cue strength through manipulating physical properties could include variations in brightness or intensity, focus or resolution, duration (through tachistoscopic means), or degree of interfering occlusions (visual "noise"). Such variations in cue intensity, with progressive diminution in the strength of the cue given, have been utilized in the SAKI device for teaching card punching described by Pask (1959b).

In programming verbal or symbolic learning, varying the effective cue strength through changes in such physical determinants of stimulus clarity might also have potential advantages, for some purposes, over the subtleties of semantic variation. For one thing, such physical variation is potentially content-free—the same mechanics of variation may be used for a large number of cueing terms and situations, rather than devised *ad hoc* for each item. For another thing, this kind of cue variation could potentially be calibrated and standardized in physical units. The latter characteristic might be of particular advantage for experimental research on cueing factors, since it could permit graded, replicated control of cue strength in different contexts.

Regardless of the means of varying the cue strength of prompts, what we should ideally achieve, for efficient learning, is an optimal *progression*, from initial maximal cueing (to guarantee initial eliciting of the desired response) down to zero or minimal cueing, as learning progresses. This progression, which Skinner (1958c) has termed "vanishing," insures that the learner is forced to obtain practice in responding with no prompt and thus does not continue to depend on the prompt. It also seems intuitively clear that this transition would be facilitated if conditions were arranged so that the learner would respond, at any point, to the *minimum cue-strength necessary at that point*. (In fact, one way of charting the course of learning would be to chart—an alternative to promptness of response or probability of its occurrence to a fixed cue situation—the strength of the prompting cue that is necessary in order to elicit the response.)

One difficulty in implementing these theoretically optimum conditions is that it is difficult to regulate or predict the minimum sufficient amount of cueing necessary at any given stage of the learner's progress. An advantage of using variation in physical cue clarity, where applicable, would be that the strength of cue can be adjusted automatically at each stage to the minimum sufficient amount required. Thus, by "fading *in*" the cue, or increasing its clarity progressively *within* a trial, the learner can respond correctly at each stage with just the barely sufficient amount of cueing. (In the later stages of learning the necessary cue intensity might be considered as being at essentially the level of the sensory threshhold.) This kind of intra-

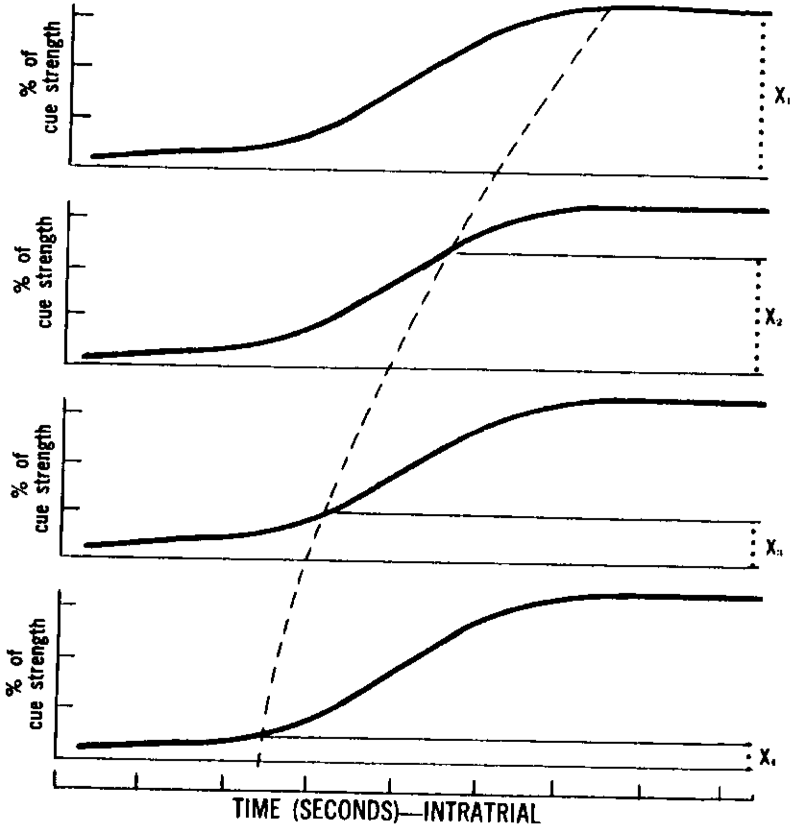


FIG. II. Hypothetical cue-strength gradients

trial build-up in cue-strength, through "fading in" the distinctiveness of the cue as the learner tries to respond, is shown in the rise in the height of each of the solid curves in Figure II.

The *increase* of cue strength *within* any trial, until it is up to the strength required to permit dependable and correct emitting of the desired response, contrasts with the progressive *decrease*, *between* trials or items, up to which the cue strength must be brought before the correct response can be emitted. This intratrial gradient is indicated by the dotted curve in Figure II. We thus have two gradients: an *intra-item* gradient of increasing cue strength (up to the point of minimum sufficiency required) and a decreasing trial-to-trial or item-to-item gradient in the stopping point on this build-up gradient for the successive trials or frames.

It is certainly not suggested that this method, despite its possible attractiveness, could be a panacea for all cueing needs: there is no thought of trying to preclude other methods of prompting, either intraframe or in preceding frames. Potentially, however, a theoretically optimal and automatically regulating device is thus provided which might be used with whatever combination of other prompting methods is desired and which is of particular interest as a research tool.

These considerations may suggest the desirability and feasibility of theoretical and experimental research which could have potentially direct practical implications for programming theory and the design of self-instructional devices, as well as more general implications for the understanding and control of human learning. The following four questions are among those to which it would be of interest to address such research.

1. (a) What cueing techniques are most appropriate for different kinds of behavioral outcomes desired, for use at different stages of learning, and for learners of varying characteristics of ability and motivation?

   (b) How should we expect their effectiveness to interact with differences in the kind of response units and response feedback that are provided for the learner?

2. In using a cue "fade-in" technique, what are the relative merits of having cue clarity governed by the program or the teaching machine (e.g., a fixed intratrial rate of increase) versus having it adjusted by the learner up to the minimum sufficient cue-strength at which he can respond correctly?

3. Can the effective cue-strength of contextual, semantic cue forms be usefully scaled against a standard of physical variation in cue-clarity or visibility?

4. What is the general validity of the hypothesis that there should be a transition from maximum to minimum cue strength? Can this principle be extended or made more precise, in terms of some quantitative parameters for optimal cue-strength gradients or supplementation in terms of auxiliary principles?

Some research pertinent to these questions has been initiated by Matthew Israel at Harvard and at the American Institute for Research with the support of the U. S. Air Force Office of Scientific Research.

A Functional Analysis of Self-Instructional Devices

W. J. CARR

THE PRIMARY purpose of this paper is to review a portion of the rapidly growing literature on the design and utility of self-instructional devices. Special attention will be given to those studies which provide for a functional analysis of such devices.

Other reviewers have concerned themselves with different aspects of the literature on automated instruction. Pressey (1958), Skinner (1958c), and Ramo (1957) have written brief histories on the subject and have discussed some of its economic, social, and even political implications. Skinner (1958c) and Gilbert (1958b, 1958c) have presented rationales for self-instructional devices that were based upon generally accepted principles of learning. Finally, Porter (1957) has reviewed the evaluation studies which compared the effectiveness of automated instruction with instruction by the more standard procedures. Porter also presented an excellent system of classification of the many types of devices which have been suggested as being useful adjuncts to the teaching process. The present paper will not be directly concerned with a review of these aspects of the literature.

Briefly, a self-instructional device might best be thought of as an automatic tutor which presents the learner with a series of problems, each of which requires some appropriate action on his part. In his review of the literature on the many types of devices employed, Porter (1957) has listed three essential features which distinguish the self-instructional device, or teaching machine, from the teaching aid. First, the self-instructional device provides a sequence of problems which are designed to take the learner from a low to a high level of proficiency in a given subject matter. Frequently, the problems also contain information which helps the learner to respond correctly. Moreover, the program of problems is graduated carefully so that the probability of responding correctly to a given problem is quite high, provided that previous problems have been correctly answered.

The second essential feature of a self-instructional device is that its use requires some action on the part of the learner at every stage of the program. With some devices, the learner must answer each problem correctly before being presented with another, while other devices permit occasional errors. The latter type usually provides for recycling of the program. The third essential feature of a teaching machine is that the device provides immediate confirmation or knowledge of results about the correctness of the learner's responses. Porter believes that devices having all three of these features might properly be called teaching devices, since no human teacher need mediate relations between the learner and the device. He suggests that other devices such as textbooks, blackboards, models, and films might best be thought of as teaching aids, since a teacher must perform one or more of the three essential functions.

PARAMETERS INFLUENCING THE EFFECTIVENESS OF SELF-INSTRUCTIONAL DEVICES

A considerable portion of the research on self-instructional devices has taken the form of evaluation studies, in which the effectiveness of teaching by machine has been compared with teaching by some more standard method of instruction. The present writer shares with Gilbert (1958b) the view that although a certain amount of evaluation research is necessary in order to justify continued interest in the basic concept of automated instruction, the major portion of research effort should be devoted to an experimental analysis of the parameters which influence the effectiveness of self-instructional devices.

A distinction may be drawn among three classes of variables of which the effectiveness of automated instruction may be a function. They are: characteristics of the device, of the pro-

---

This paper was prepared under the auspices of the Aerospace Medical Laboratory of the Wright Air Development Center of the U.S. Air Force. A substantially identical version has recently been published by the Air Research and Development Command as WADC Technical Report 59-503. Dr. Carr is now at Temple University, Philadelphia, Pennsylvania.

gram, and of the learner. Each of these classes of variables will be discussed in some detail, and an effort will be made to summarize the pertinent research findings.

CHARACTERISTICS OF THE DEVICE

For purposes of discussion, a self-instructional device may be thought of as having four major components: (a) a display by which the program is presented, (b) a response panel which the learner uses in forming his response, (c) a confirming mechanism which provides the learner with information as to the correctness of his response, and (d) a reinforcement mechanism which provides the impetus for further operation of the device.[1] This fourfold analysis establishes a convenient framework in which to place a discussion of research on characteristics of self-instructional devices.

Display or Input Characteristics

"Input-throughput-Output" ?

Porter (1957) has reviewed the literature on the devices which qualify as self-instructional devices, according to his criteria. No attempt will be made to describe the devices in the present paper. Suffice it to say that workers have used mechanical, electrical, and optical components to present materials to the learner, usually by visual or auditory stimuli.

Moreover, Porter has distinguished between two important ways of presenting the program. The first way might be called "learner-paced," in that the device waits upon the learner to respond before it reacts. Both Pressey and Skinner have made use of this technique. The second way of presenting the program might be called "machine-paced," in that a given problem is presented for a period of time and then the machine reacts whether or not the learner has responded to the problem. The well-known memory drum is an example of this latter way of presenting the program. Finally, one purely mechanical device seems sufficiently unique to warrant special note. Crowder (1959) and Homme and Glaser (1959b) have developed successful programs in the form of a radically modified textbook.

---

[1] The confirming and reinforcing mechanisms are assumed to be combined in some types of self-instructional devices.

Gilbert (1958c) has discussed some of the practical problems connected with the presentation of materials, but to the writer's knowledge, no investigator has compared the effectiveness of two or more methods of presenting the program to the learner. One research problem which might well receive immediate attention is the relative efficiency of group-paced versus learner-paced devices. Lumsdaine (1959a) has pointed to some of the similarities and differences between the two types of presentation.

Obviously, a great deal of the human engineering type of research is necessary before we know much about the optimum means of presenting materials for automated instruction. Furthermore, it is likely that the most effective means of presenting the program will vary as a function of such other variables as the response characteristics of the device, the type and amount of subject matter to be learned, and perhaps with characteristics of the learner.

Response or Output Characteristics

Two basically different modes of operating teaching machines have been employed. Following the lead of Pressey (1950), many have used the recognition, or multiple-choice, method of responding. Others have employed Skinner's recommended construction, or fill-in, method of responding. Gilbert (1958c) has discussed some of the practical problems connected with the two methods of responding. [See Coulson and Silberman (1959b) and Fry (1960a). —Eds.]

Evans, Glaser, and Homme (1959) have compared the effectiveness of the construction method of responding with no overt responding at all. In this experiment, two groups of *Ss* learned a program entitled "fundamentals of music." One group made one or more written responses to each problem, while the other group made no overt responses. Interestingly enough, the group which made no overt responses spent less time in learning and showed slightly higher retention test scores than did the group which responded overtly to each problem; however, the differences were not statistically significant. Obviously, this variable is worthy of more careful attention, since it is possible that at least some responses need only be made covertly.

Characteristics of the Confirming Mechanism

Gilbert (1958c) has defined the confirming mechanism of a teaching device as the means by which the learner receives information as to the correctness of his response to a given problem. The designers of self-instructional devices frequently assume that confirmation also provides reinforcement. However, the present writer agrees with Gilbert, who implies that confirming and reinforcing mechanisms might best be discussed separately, since they refer to separate classes of independent variables, of which the effectiveness of automated instruction might be a function.

Confirming mechanisms employed in self-instructional devices can be varied in at least two important dimensions. First, they might vary in the type or amount of information provided to the learner, and, second, they may vary in the temporal delay between the response of the learner and the delivery of the confirmation. These subclasses of variables will be discussed separately.

Irion and Briggs (1957) have described four modes of confirmation appropriate to the Pressey type of self-instructional device . . . . [The four confirming arrangements described by Irion and Briggs are: (a) *quiz mode,* in which the machine immediately indicates which of a number of response options is correct; (b) *modified-quiz mode,* in which the learner chooses an option, being presented at once with the next question if correct; if incorrect, an error signal is accompanied with a direct prompt that shows which option he must then choose to be correct; (c) *practice mode,* in which the learner is required to keep choosing options until the correct one is chosen; and (d) *single-try mode,* in which the learner can respond only once to each problem; a light indicates whether the response is correct, after which the next problem is presented. The first two of these four modes correspond, respectively, to those described on page 300 by Briggs (1958) as the "coaching mode," and the "single-error-permitted" mode.—Eds.].

Irion and Briggs compared the effectiveness of the four types of confirming mechanisms on each of three types of learning tasks: serial learning, paired-associate learning, and problem solving. Independent groups of 20 subjects each were given 20 minutes of practice under one of the 12 experimental conditions. The dependent variable in the experiment was the number of errors committed on a retention test given immediately after practice and two weeks later. In general, the relative efficiency of the four types of confirming mechanisms was in this order: quiz mode, modified-quiz mode, practice mode, and the single-try mode. The amount of difference in efficiency varied as a function of the type of learning task.

Unfortunately, the types of confirmation studied by Irion and Briggs were such that a number of more basic independent variables probably were confounded. For example, as Irion and Briggs point out, it may be that the quiz mode of confirmation proved to be generally the most effective because it permitted more practice trials per unit time, because it reduced the number of errors which could be committed, because it eliminated the aversive stimulation resulting from the error buzzer, or because of some combination of these variables. Although Irion and Briggs answered an important practical question, it would seem that ultimately the problem of optimizing the confirming mechanism of a self-instructional device must be attacked by dealing with but one variable at a time.

Investigators who prefer to have the learner construct his answers to problems rather than simply recognize them have also employed at least two types of confirming mechanisms. Skinner (1958c) has used a device which presents the learner with the correct answer to a problem as soon as the learner completes his response, whether or not the response is correct. By recycling the program, the learner eventually answers all problems correctly. Recently, Skinner (1959b) expressed dissatisfaction with this design of machine, because the second attempt to answer a given problem may be partly under the control of the previous revealed response. Skinner has also developed a device which continues to present a given problem until the learner responds correctly. These two types of confirming mechanisms parallel Irion and Briggs's quiz mode and practice mode respectively.

Gilbert (1958c) has discussed some of the practical problems connected with the use of the quiz mode and practice mode of responding, and he has suggested a third method of responding. It incorporates features of both the construction and the recognition methods of responding. When faced with a given problem, the learner first constructs his answer, then he views a number of alternatives and chooses the one which best approximates his response. The machine then informs the learner in respect to the correctness of his choice. Gilbert has also suggested that the confirming mechanism might also present additional information at the same time that it confirms the learner's response. The promise of this latter suggestion finds support in the findings of Evans, Glaser, and Homme (1959) who showed that the learner need not respond overtly to all of the materials in a program. To date, none of these suggestions have been subjected to experimental comparisons using the construction method of responding.

The second dimension along which the confirming mechanism of a self-instructional device might vary is the temporal delay between the learner's response and the delivery of confirmation. It is frequently stated that one of the chief advantages of automated instruction over the more standard instructional procedures stems from the immediacy of confirmation. However, Porter (1957) has reviewed the literature on such evaluation studies, and he reports that in most experiments the subjects who learned by machine probably received more confirmation as well as receiving it more rapidly. Thus we are left without knowing much about the separate roles played by these two variables.

One must go outside of the context of research on automated instruction to gain support for the belief that the rate of learning is a function of the temporal delay between the learner's responses and the confirmation. Saltzman (1951) had two groups of *Ss* learn a list of six paired-associates. One group received immediate confirmation. For the other group, a six-second delay intervened between the disappearance of the stimulus word and the appearance of the correct response. Saltzman reported a 50 percent increase in trials to criterion for the group which received the delayed confirmation. It is possible that increasing the amount of material to be learned might appreciably increase the effect of delay of confirmation upon learning time.

Additional research on relations between delay of confirmation and learning time seems badly needed, especially since certain types of electronic computers have been considered as self-instructional devices (Rath and others, 1959). The modification of such devices may render them capable of instructing simultaneously large numbers of learners. Each learner might have his own display and response panel which could be connected with a central computer-like device. The latter would both present the program and react to the responses of many learners at one time. Obviously, the maximum number of students who could use such a device would depend upon the maximum permissible delay between the responses and confirmation.

The preceding discussion of the confirming mechanism used in self-instructional devices implies that learning is most rapid if confirmation follows the learner's attempt to answer a given problem. However, some recent evidence suggests that a technique called *response prompting* might prove even more effective (Cook and Kendler, 1956; Cook, 1958; Cook and Spitzer, 1960). These investigators compared the effectiveness of two procedures in a paired-associates learning situation. The confirmation technique involved the presentation of the response term after the learner had attempted a response. The prompting technique involved the presentation of the response term before the learner made his response. The results of their experiments showed that the response-prompting technique yielded faster learning than did the confirmation technique. Moreover, similar results were obtained by Kopstein and Roshal (1955) who also reported that the advantage of response prompting over confirmation decreased as the degree of learning increased. Obviously, further research of the sort here described is very much needed.

Reinforcement Characteristics

The writer has distinguished between confirming and reinforcement mechanisms inherent in teaching machines on the grounds that they refer to two distinct classes of variables which probably influence the effectiveness of automated instruction techniques. Confirming the correctness of a learner's responses to problems may be expected to be reinforcing only if the learner's motivation is intrinsic to the task being learned (Gagné and Bolles, 1959).

Consideration should be given to a number of questions having to do with the reinforcing characteristics of self-instructional devices. First, it is frequently stated that perhaps the most important benefit resulting from automated instruction is the frequent reinforcement which the learner receives. However, an important question to answer is whether the benefit results from a practice effect or from an increase in motivation to learn. No experiments which bear directly on this problem have been uncovered. However, Michael and Maccoby (1953) performed a relevant experiment on the effect of training films upon test performance. These workers conclude that the benefit from audience participation seems to stem primarily from the effects of practice and not from an increase in motivation. More research seems required on this important question within the context of automated instruction.

A second question having to do with the reinforcing characteristics of self-instructional devices stems from a warning made by both Porter (1957) and Keislar (1959a) of a *novelty effect* in connection with the use of confirmation as the sole reinforcement. It may well be that confirmation would lose its reinforcing property over time. Porter (1958c) reports that the advantage of automated instruction over a standard method of teaching showed no diminution over a period of five months. However, more research of this sort is necessary before we can be sure that confirmation can be relied upon to maintain its reinforcing property for very long periods of time.

A third question has to do with the optimal schedule of reinforcement to employ. Most researchers agree that the program should be carefully calibrated, so that the probability of the learner's answering questions correctly is very high. Such a situation would result in a schedule which approaches 100 percent reinforcement. Skinner and Holland (1958) have implied that the completion of a given number of problems also constitutes reinforcement using his machine. Thus he feels that learners are operating under a special type of partial reinforcement schedule which has been called reinforcement on a fixed ratio with counter. Further, both Pressey (1950) and Skinner (1954) feel that extrinsic reinforcement might also be employed. No research has been uncovered which tells us about the optimal kinds and schedule of reinforcement to use.

Finally, one should be careful to distinguish between two kinds of events which might get reinforced. The first event is responding, per se, and the second is responding correctly. Presumably, the second kind of event is the proper one to reinforce. Skinner also reports that learners have a tendency to respond carelessly during the first cycle of the program, trusting to a look at the correct answer to clear up the assignment. He attributes this tendency to too high a level of motivation stemming from the design of his machine (presumably the reinforcing mechanism), and he has suggested a technique to circumvent the problem (Skinner and Holland, 1958).

It seems clear from this brief discussion that the nature of the reinforcement employed in self-instructional devices has not yet been carefully analyzed, let alone adequately studied.

CHARACTERISTICS OF THE PROGRAM

Programming refers to the arrangement of materials to be learned in that order of presentation which will maximize the rate of acquisition and retention. It is possible that some simple tasks require little in the way of programming, but it is very likely that the development of efficient programs to be used in the teaching of complex skills must wait upon a thorough functional analysis of the characteristics of a program that influence both rate of acquisition and retention.

Attempts have been made to program many different kinds of subject matter, ranging from the teaching of contract bridge (Crowder, 1959) to the development of managerial skills (Pask, 1957a). However, the present paper is not concerned with the programming of any particular kind of subject matter, but rather with the characteristics which are probably common to many kinds of materials.

Skinner (1954) was the first investigator to give serious attention to the problem of programming for automated instruction. More recently, Gilbert, (1958b, 1958c) has described in considerable detail some of the principles of program composition. In addition, Crowder (1955, 1958c, 1959), Meyer (1959), Beck (1959), and Glaser, Homme, and Evans (1959) have made important contributions to our understanding of the process. The description which follows is an attempt to synthesize some of the views of these investigators. The analysis is directly concerned with the development of verbal and symbolic skills, but it might also apply to the development of psychomotor skills.

Meyer (1959) has described three major steps involved in composing a program for the automatic teaching of a given subject matter. A description of these steps may be couched either in terms meaningful to the curriculum specialist or to the psychologist. In the interest of clarity, the writer intends to use both levels of description.

Programming as the Curriculum Specialist Views It

From the point of view of the curriculum specialist, the first step in program composing is to delineate the entire field or subject matter to be taught. Terms, methods, facts, principles, theories, etc., must be collected. One might think of these as answers to questions which might appear on a final examination for a given course. The second step is to ascertain the learner's level of understanding of the subject matter before any training has begun. The third step is to arrange the subject matter into a logical order which is conducive to rapid learning and good retention.

The arrangement of the subject matter into a logical order can be further analyzed into three distinct steps. First, a hierarchy of the materials must be established so that the learner will first master elementary skills which he will later use to develop more complex ones. Second, the hierarchy of materials must be composed of steps small enough to be taken readily by the learner without being so small as to impede learning. Finally, the program must provide for sufficient conditioning at each step in order to be sure that each step will be adequately learned. Otherwise, extinction might well take place before the skill is put to use later in the program.

Programming as the Psychologist Views It

The psychologist who is interested in the study of the process of learning might describe in quite different terms the steps involved in composing a program for use in self-instructional devices. Presumably, the psychologist's major contribution to the study of automated instruction techniques would stem from the understanding of the optimal means of effecting the transition from low to high levels of proficiency with respect to the subject matter to be learned. For this reason, a somewhat detailed analysis of the program seems required in an effort to isolate the variables of which the effectiveness of a program might be a function.

For ease of communications, one might characterize an entire program as consisting of three types of stimulus-response connections.[2] Let *terminal S-R connections* refer to the stimulus-response connection which we desire the student to learn. Let *initial S-R connections* refer to those stimuli which at the outset of training are already the occasions for those responses which, to some degree, approximate the responses the programmer desires to teach. Finally, let *transitional S-R responses* stand for the steps which can be expected to mediate the initial and terminal S-R connections. A number of transitional S-R connections might be required to bridge the gap between a particular pair of initial and terminal S-R connections. The first transitional S-R connections would only crudely approximate the terminal connections, both with respect to stimuli and responses. The second transitional connections would approximate the terminal connections more closely, the third even more, and so until the terminal S-R connections are reached. Figure I shows a schematic drawing of this conception of a program.

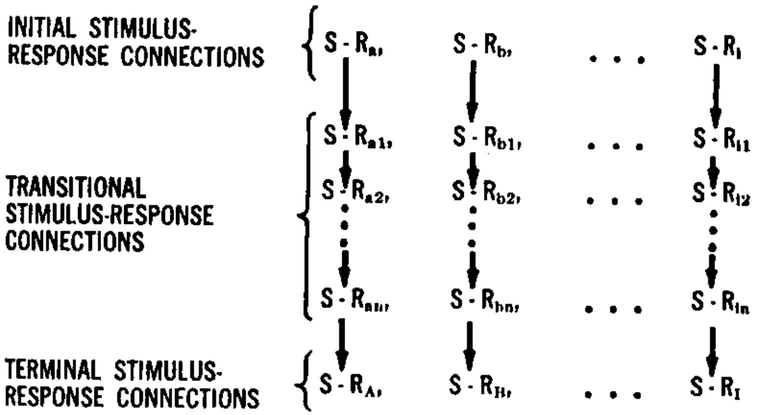

INITIAL STIMULUS-RESPONSE CONNECTIONS
$S-R_a$, $S-R_b$, . . . $S-R_i$
TRANSITIONAL STIMULUS-RESPONSE CONNECTIONS
$S-R_{a1}$, $S-R_{b1}$, . . . $S-R_{i1}$
$S-R_{a2}$, $S-R_{b2}$, . . . $S-R_{i2}$
$S-R_{an}$, $S-R_{bn}$, . . . $S-R_{in}$
TERMINAL STIMULUS-RESPONSE CONNECTIONS
$S-R_A$, $S-R_B$, . . . $S-R_I$

FIG. I. Matrix showing the development of a program from initial stimulus-response connections, through the transitional connections, and finally approaching the terminal stimulus-response connections.

---

[2] In this paper the term *stimulus-response connection* is used only as a matter of convenience. The writer does not imply an absence of covert stimuli and/or responses which mediate the connection.

The psychologist might continue his analysis of programming by pointing to two problems connected with the sequencing of materials to be used in automated instruction. First, the learner must be caused to emit the appropriate responses, and, second, these responses must be brought under the control of the appropriate stimuli. Clearly, these operations cannot be performed in one great step from the initial to the terminal S-R connections. Rather, a series of intermediate steps must be established which guide the learner at every point along the way. At the outset the program should consist primarily of initial S-R connections. The topography of these connections must undergo a systematic change to become transitional S-R connections and these must undergo changes until the terminal S-R connections are reached.

Various investigators have described three steps involved in guiding the learner through the program. In order to simplify matters, consider the development of only one terminal S-R connection. First, the stimulus of the initial S-R connection is paired with the stimulus of the first transitional S-R connection. The reader will recall that the former type of stimulus is, by definition, already the occasion for the response which approximates the desired one. When these two kinds of stimuli are paired, the learner's first response should be the response of the initial S-R connection. Gilbert (1958b) calls this process *augmenting*.

The second step in the process is to extinguish the response of the initial S-R connection through nonreinforcement,[3] thus permitting the first transitional response to appear and be strengthened through reinforcement (Skinner, 1958c). The third step in the process is gradually to eliminate the stimulus from the initial S-R connection, leaving only the stimulus from the first transitional S-R connection. This process has been called *fading*, or *vanishing* (Skinner, 1958c; Gilbert, 1958b). The connection is said to be established when only the stimulus from the first transitional S-R connection evokes the response from that connection.

The three steps may then be repeated by pairing the stimulus from the first transitional S-R connection with the stimulus from the second transitional S-R connection. Then, by differential

---

[3] In this case, extinction is assumed to be either stimulus or response produced, and, in either case, it does not necessarily require overt responses on the part of the learner.

reinforcement, the response of the first transitional connection is weakened and the response of the second transitional connection is strengthened. Next, fade or vanish the stimulus from the first transitional connection, leaving only the second, until the second transitional connection is firmly established. This process is repeated over and over until finally the terminal S-R connection is established.

At least two classes of independent variables are suggested by the present analysis. The first variable has to do with the amount of stimulus augmenting built into the program. Put differently, the extent to which the correct response is prompted by stimuli which accompany the problem is a variable of which the efficiency of a program might be a function. The implication is that there is an optimum degree to which the stimulus of the first transitional S-R connection should be augmented by the stimulus of the initial S-R connection. Too high a degree of augmenting would be wasteful of learning time, while too low a degree of augmenting would result in frequent errors.

The second variable suggested by the present analysis has to do with the rate of fading of the augmenting stimuli, once the response of the first transitional S-R connection has been evoked. Once again, there should be an optimum rate of fading, or vanishing, of these augmenting stimuli. Too slow a rate of fading would be wasteful of learning time, and too rapid a rate would result in frequent errors.

The reader will note that the frequency of errors committed by the learner as he proceeds through the program can be manipulated either by varying the amount of augmenting or the rate of fading. The effect of frequency of errors upon the efficiency of a program will be discussed later in this paper.

In addition to the methods already described, Skinner (1958c), Beck (1959), and others have suggested a number of techniques which might be used to evoke correct responses from the learner at every point in the program. Beck has called these *rules of programming* and has attempted a classification of them. One technique consists of giving the learner some materials to read either before he sets to work on the program or while he is working on it. Another technique involves the use of the context of the problems as a means of eliminating some strong competing response. For this purpose, the correct answer might require a word which rhymes with a word in the problem or which is the opposite of a word in the problem. All sorts of techniques have been suggested. However, it seems clear that these techniques simply provide stimuli which reduce the probability of an error response which might otherwise occur. In a sense, the techniques provide for something akin to stimulus-produced extinction.

It is clear that the conception of a program described in the present paper does violence to programs as they appear in real life. Skinner (1959b) has already noted that it is probably unwise to set about teaching students specific responses to specific stimuli. We want students to learn to respond correctly to questions about a particular subject matter, but we also want them to be able to deal effectively with related materials not presented in the program, and, certainly, we don't want them simply to parrot responses as they appear in the program. We want them to give answers in their own words. The writer would agree that it is probably impossible ever to specify completely the terminal S-R connections in a given subject matter, and this probably also applies to the initial and transitional S-R connections as well.

However, such considerations should not lead us to abandon model building, especially if the model permits us to see more clearly some of the variables which influence behavior in a learning situation. The present analysis says nothing new. Researchers interested in programming have implied all of these processes. The present writer simply wishes to make the processes more explicit.

It should be noted that the matrix shown in Figure I is defective on at least three additional counts. First, the matrix fails to show that some transitional S-R connections are also terminal S-R connections. Second, it fails to make clear that two or more transitional connections might be combined to form but one terminal connection and, third, the matrix fails to show that a given transitional connection might play a role in the formation of more than one terminal connection. These three defects are viewed by the writer as trivial; a more complex matrix could incorporate all of these features.

Sources of Error in Program Writing

It is relatively easy to describe in generalized stimulus-response terms the nature of a program to be used for the automatic teaching of a subject, but it is quite another matter to specify precisely how to go about composing one. At least five sources of error may enter the process. First, the programmer may incorrectly specify the sum total of the terminal S-R connections to be formed. This would amount to saying that the curriculum specialist failed to delineate completely the subject matter to be taught. Second, the programmer may err in his estimate of the extent of the initial S-R connections, which amounts to saying that the curriculum specialist overestimated or underestimated the learner's level of understanding of subject matter before training has begun. Third, the programmer may not provide for sufficient conditioning of one or more transitional or terminal S-R connections. This error would be comparable to the failure on the part of the curriculum specialist to provide for sufficient practice on any aspect of the subject matter. Fourth, there may be one or more defects in the order of the progression of transitional S-R connections. In the parlance of the curriculum specialist, such a defect would amount to teaching complex skills before the learner has mastered more elementary skills which make up the complex ones. Finally, the programmer may make premature progressions in the topography of the transitional S-R connections. To the curriculum specialist, this means that the level of complexity of the subject matter is being raised too rapidly and the learner cannot keep up.

In considering these five sources of error inherent in the process of composing a program, one should keep in mind the important concept of individual differences. For example, the initial S-R connections and the amount of conditioning required at each transitional step may differ considerably from one learner to another. This problem and ways of dealing with it will be examined later in this paper.

Any of these five sources of error may seriously impair the effectiveness of a program to be used for automated instructional purposes, and the sheer number of sources may lead some to believe that there is little chance of writing successful programs. Indeed, some of these errors have already been reported in the literature. Meyer (1959) apparently overestimated the extent of the initial S-R connections. In developing a program for the teaching of arithmetic to elementary-school children, she assumed that the children could match numbers which were presented as stimuli with numbers to be used as responses. The program also assumed a minimal reading vocabulary (such words as "is," "here," "and," etc.). Later Meyer found that these assumptions were not warranted.

Moreover, Keislar (1959a), who developed a program for the teaching of certain mathematical concepts, also apparently erred, either by making premature transitions of the transitional S-R connections toward the terminal connections or by failing to provide for sufficient conditioning once the connections had been established. Keislar reported that subjects made too many errors in going through the relatively short program. Undoubtedly, other investigators have committed one or more of the five kinds of error here described without reporting this. It is interesting to note that both Meyer and Keislar attributed the faulty performance of the learners to defective programs. Although such errors clearly point to the need for further research, it should be noted that teachers who employ the more conventional methods of instruction might be less inclined to be so self-critical.

In any event, the sources of error which enter into the composing of a program to be used in a teaching device are precisely the same ones which enter into the use of standard pedagogic methods. Further, although the evidence is far from clear-cut, it appears that when self-instructional devices are experimentally pitted against standard methods of instruction, the former prove to be the more effective (Porter, 1959; Evans, Glaser, and Homme, 1959).

A FUNCTIONAL ANALYSIS OF THE PROCESS OF PROGRAMMING

Several investigators have described what have been called principles of programming; and although there has been considerable overlapping, each worker has contributed something new. The present writer has attempted to incorporate into one statement the principles which have been suggested by Skinner (1954), Gilbert (1958b, 1958c), Meyer (1959), and Glaser, Homme, and Evans (1959).

However, it should be noted that these principles of programming, as stated by most workers, simply constitute problems which the programmer faces when he attempts to compose a program. As such, they offer no solutions, and presently each programmer must solve these problems as best he can. In the interest of scientific efficiency, each of these principles might better be thought of as a class of independent variables, of which the efficiency of a program might be a function. Glaser, Homme, and Evans (1959) appear to think of them this way, and these investigators have already begun a functional analysis of one class of variables and have suggested a line of attack upon another. A functional analysis of the process of programming provides a convenient framework in which to summarize previous research findings and to suggest new approaches which might be taken.

1. Relevance

At the outset, the programmer must specify precisely the terminal S-R connections to be formed, i.e., what responses are to be brought under the control of what stimuli. Skinner (1954) first described this principle, but the term *relevance* comes from Gilbert (1958b) who stated that the problems to be presented in the program should represent the kinds of problems which you want the learner to solve and the responses in the program should approximate the responses you want him ultimately to make.

Glaser, Homme, and Evans (1959) have reference to relevance variables when they speak of behavioral end-products, and they emphasize the importance of specifying precisely what form the skills to be learned are to take. For example, the optimal properties of both the teaching device and the program are likely to depend on whether the learner is to acquire facts, solve problems, or make practical applications of the materials to be learned. This is the old problem of transfer from the learning situation to the task for which the learning is intended. Porter (1957), Kendler (1959), and Keislar (1959a) have discussed some of the problems of transfer which might arise from the use of self-instructional devices. In short, these investigators seem to be saying that a complete functional analysis of the process of programming would include consideration of the problem of transfer from the training situation to the actual task situation.

2. Availability

The programmer must also specify precisely the initial S-R connections, i.e., those connections already in the learner's repertory which approximate the terminal S-R connections and from which the transitional S-R connections are to be developed. Skinner (1954) and Meyer (1959) have described this principle, but it seems to have been overlooked by others.

To the writer's knowledge, no research has been done on the problem of specifying the initial S-R connections on which the program is to be built. Presumably, most researchers think of this as a problem for the curriculum specialist or the psychological tester.

3. Sequencing

The programmer must specify the optimum order of presentation of the transitional S-R connections which will enable the learner to proceed from the initial to the terminal S-R connections in such a way as to maximize the learner's performance on some criterial measures of learning and retention. Gilbert (1958b) described this class of variables, but he had reference to the sequencing of machine instruction with expository teaching and simulated field experience. In the present paper, this class of variables is given a broader definition so as to include both the views of Gilbert and those of Glaser, Homme, and Evans (1959).

Although no research has yet been done within the context of automated instruction on the problem of how best to sequence a program, two suggestions have been made. Skinner (1958c) feels that if possible, a single program should be developed through which all learners must proceed. This might be thought of as a straight-line or linear program. On the other hand, Crowder (1955, 1958c, 1959) has developed a program which permits the branching of the subject matter so that allowances might be made for characteristics of the individual learner. Crowder calls his type an intrinsic program and reserves the term extrinsic for the type Skinner has described. More recently, Gilbert (1958c) has presented a rationale which leads him to prefer the extrinsic over the intrinsic type of program, but he points out that the issue must ultimately be resolved by the appropriate experiments.

Whether one uses the extrinsic or the intrinsic type of program, the problem of sequencing still remains. Proper sequencing would permit the most rapid shaping of the learner's behavior and result in maximal retention. Earlier in this paper the process of shaping has been described in terms of the well-known processes of stimulus discrimination and response differentiation, and the writer has proposed two independent variables of which the efficiency of a program might be a function. So far, no research of this type has been published.

4. Stepping

After the proper sequence of transitional stimulus-response connections has been developed, the programmer must specify the size of a step from one transitional S-R connection to the next. The size of step can be defined operationally in at least two ways. When it is used as an independent variable in an experiment, it is usually specified as the number of steps in a program which takes the learner from the initial to the terminal stimulus-response connections (Homme and Glaser, 1959b). The greater the number of steps, the smaller the median size of step. When the term is used as a dependent variable, it is usually specified by the percentage of incorrect responses. Thus if learners make few error responses on a given program, the size of step is inferred to be small.

At one time Skinner (1958c) believed that the size of step should be so small that the learner rarely if ever makes error responses, but more recently (Skinner, 1959b), he has adopted the position that the optimal size of step is an empirical question, presumably to be answered by finding that size of step which maximizes learning and retention.

In one of the first experiments designed to provide a functional analysis of the process of programming, Evans, Glaser, and Homme (1959) investigated the effect of number of steps in a program upon learning time, the frequency of errors during learning and upon immediate and delayed test performance [see pages 447 to 448]. The results show that, within limits, increasing the number of steps in the program resulted in decreases in the number of errors on immediate and delayed performance tests. In addition, smaller steps also resulted in less time per step and fewer errors during the course of learning.

Evans, Glaser, and Homme point out that the optimum size of step might be expected to vary as a function of the type of subject matter being programmed. Clearly, further research is in order before this important question can be answered unequivocally, but a start has been made.

5. Maintenance

The programmer must specify the amount of conditioning required of both transitional and terminal stimulus-response connections to guarantee adequate learning and maintenance of these connections. Gilbert (1958b) has discussed this problem in some detail, under the principle of repetition. He states that optimum repetition required the use of a minimum number of problems and a minimum sample of problems in a given class and a minimum amount of time invested per student which will produce a satisfactory probability of correct responses. He suggests that review materials be seeded at various points in the programs to be sure that transitional stimulus-response connections will be maintained. Although the problem of repetition or maintenance lies clearly within the province of the psychologist interested in the learning process, no experiments were found on this problem as it relates to automated instruction. However, the abundant data on overlearning and underlearning (Woodworth and Schlosberg, 1954: 728-32) seem appropriate.

CHARACTERISTICS OF THE LEARNER

A third major variable of which the effectiveness of automated instruction might be a function has to do with characteristics of the learner. The concept of individual differences has been implicit throughout the preceding discussion of the variables which might be expected to influence the efficiency of a program to be used for the automatic teaching of a given subject matter. For example, the kind and number of initial stimulus-response connections available to the programmer which he can use as starting points for the program will obviously depend upon the learner's previous reinforcement history. Moreover, the learner's intelligence and his aptitudes and interests with respect to the subject matter being taught might influence the characteristics of the program having to do with repetition, sequencing, and stepping. Several workers have implied that the concept of individual differences is important to a discussion of the variables which affect program efficiency—e.g., Glaser, Homme, and Evans (1959).

One might hypothesize that effective instructional devices might wipe out differences in achievement measures associated with intelligence of aptitude test performance. The findings of a number of experiments seem to support this hypothesis. Porter (1959) found that the correlation between IQ and achievement in spelling was not significantly different from zero for a group of learners taught by a self-instructional device, but among the control subjects taught in the standard fashion a significant positive relationship was found. Irion and Briggs (1957) reported "little relationship" between intelligence (Otis) and retention after learning by self-instructional devices. Ferster and Sapon (1958) reported similar findings between aptitude and achievement in a course in German. Finally, Hovland, Lumsdaine, and Sheffield (1949) reported that the advantage of active or passive participation in a simple learning task decreased with increases in intelligence.

One explanation which might be offered to account for the decrease in correlation between achievement and intelligence or aptitude test performance is that the learning of a subject matter by teaching machine renders the learners more homogeneous with respect to achievement scores. Homme and Glaser (1959b) and Evans, Glaser, and Homme (1959) report data in support of this hypothesis, but Keislar (1959a) found that machine instruction rendered the learners more variable on the achievement measure than the control group. Clearly, more research of this sort is necessary before we can be sure of the effects of automated instruction upon relations between achievement and its classical predictors.

Porter (1959) is the only investigator who has reported on relations between a few nonintellectual factors and achievement test scores earned by learners who used the teaching machine. He found no relationship between the sex of the student or the liking of the instructional method and achievement. These and other nonintellectual factors, such as level of anxiety of the learner, deserve careful attention.

It should be noted that the studies described in this section simply related some characteristics of the learner to achievement resulting from automated instruction as opposed to instruction by way of more standard teaching methods. Although such studies answer an important question, they do not get at many possible interactions between characteristics of the learner and the other variables which influence the effectiveness of automated instruction. For example, a study relating optimum size of step to the intelligence of the learner seems very much in order. Such a study and others like it would add the concept of individual differences to the other classes of variables, all of which must be examined as a part of a functional analysis of automated instruction.

Concluding Remarks

A TECHNOLOGY OF INSTRUCTION

In preparing this volume the Editors have had as their primary objective the collecting and careful editing of a number of papers in order to bring together in one place the somewhat scattered beginnings of a new direction in teaching and learning. It seems to us that the numerous contributors whose writings have produced this volume have reflected one dominant idea. This is the concept that the processes of teaching and learning can be made an explicit subject matter for scientific study, on the basis of which a technology of instruction can be developed. Despite some differences in approach and focus of interest, this underlying unity of conception and purpose emerges clearly from these diverse beginnings. More important than the role of devices and programs themselves is the fact that a conscious attempt to utilize these means of instruction forces the educator to become explicit both about his instructional goals and about the processes he employs as means of attaining them.

The papers collected in this volume represent only the *beginning* of the effort that is needed to implement the consequences of this point of view and to develop the technology of instructional method which we believe it can generate. The thinking and experimental endeavor that went into the production of these papers should in turn facilitate further investigation and discovery. At this early stage it seems premature to predict the specific outcomes toward which sustained effort in this direction may lead. What does seem predictable is that the explicit identification of variables and functional relationships which influence the modification of behavior through instruction will continue to be the heart of effective attempts to improve education. Application of principles thus discovered is the basis on which the systematic development of teaching machines and programmed learning must proceed.

Much of the present work may be transient in its specific significance. As research and development proceed, the focus of attention might shift from the specific variables and instructional techniques which have been described by the writers of these papers. As such, the terms "teaching machines" and "programmed learning" might disappear, perhaps to be replaced by some more generic term such as "auto-instructional methods." Future experimentation may alter current ideas about transfer to criterion performance from constructed or selective responses. The role of reinforcement and feedback may be differently conceived, and present conceptions of the importance of overt response might change. Emphasis on fixed- versus variable-sequence programs may also be changed, both by the discoveries of experimental psychologists and by engineering innovations which affect the relative feasibility of implementing different approaches. Current approaches to program construction might give way to new methodologies. For example, the possible emergence of economically feasible computerized systems might modify present notions of the process of programming.

Any or all of these and other changes may take place, and some of them doubtless will, without altering the basic conception that instruction and learning are amenable to systematic description and improvement through experimental inquiry. The significance of the approach represented by teaching machines and programmed learning lies primarily in its assumption that this is possible, rather than in the specific characteristic of machines, of programmed materials as they currently exist, or of theoretical points of view that underlie current efforts.

As we learn more about learning, teaching can become more and more an explicit technology which can itself be definitively taught. The belief that teaching is primarily an art with which the gifted teacher has to be born and which defies precise description thus gives way to the conviction that teaching consists of techniques and procedures which can, in large part, be made communicable or teachable. This is not to say that the talent of the superior teacher can be replaced. On the contrary, it seems clear that outstanding performance in teaching, as in any profession, is achieved only by those who, in addition to a firm grounding in a communicable technology, bring to their practice a high degree of creativity and inspiration. This certainly must remain true in teaching as well as in medicine, law, architecture, engineering, physics, or musicianship. At the same time, however, the highest achievements in any profession seem likely to be realized only when they build upon a well-developed underlying technology.

RESEARCH, DEVELOPMENT, AND PRODUCTION

For the successful development of such a technology, concerted attention should be given both to attempts at applied development and to careful, analytic, supporting research. The venturesome tryout of new approaches to the use of programmed instructional material in our schools can teach us a great deal. The reciprocal interaction between laboratory research and practical tryout in the classroom seems bound to be fruitful. Hypotheses for research and suggestions for practical innovation each stem both from theoretical considerations and from discoveries made in the course of attempts at application.

The need for flexibility and growth is an obvious one in any infant endeavor. It is important to avoid any tendency to "freeze" either the design of teaching machines or the methodology of program construction. Innovations in methods of programming and the results of experimentation will lead to the need for revision in current designs and program formats. For example, experimentation with various forms of branching, response mode, cueing, and feedback to the learner may lead to new requirements for device capabilities.

At the present time many individuals are anticipating "the coming boom in teaching machines." One of the consequences of a tendency to "get into production" is the adoption of a design for machines prior to the availability of suitable programs or of adequate information about program characteristics. If such machines are purchased in large quantities, there may be a tendency to fix the characteristics of the programs that can be used with the particular machine. This seems to us to be putting the cart before the horse. The development and production of a variety of machines is to be encouraged for experimental use and developmental tryout; however, the production of a machine needs to be accompanied by the concurrent production of a sufficient number of tested programs, suitable for use with it, so that an adequate test can be made of its utility. It is obvious that the machine cannot be tested as such, but can be tested only as a vehicle for the type of programmed sequence that it can present.

In addition to the production of a reasonable number of developmental models for use in school systems, there is also the need for a wide variety of experimental devices for research purposes. Here, however, the economic constraints of the practical teaching situation do not apply. If an experimental device has characteristics which will lead to new knowledge about auto-instructional techniques, it may represent a worthwhile investment of resources even though it is too expensive, too bulky, or otherwise impractical for wide-scale use in school systems.

In the production of programs a major problem could arise from premature publication and sale of hastily conceived and untested programs. This would pose a serious problem, since flooding the market with such programs could tend to discredit the whole concept of programmed introduction. It would therefore appear that a high-priority objective is that of working out acceptable quality-control standards for programs. Such standards are more applicable to the development of programs than to specific devices, since it is the programs that carry the burden of subject-matter content and the sequence by which it is taught. In the development of quality-control standards for programs, it is important to avoid the imposition of inflexible requirements which might inhibit creativity and experimental use of new techniques.

It seems clear that standards for the adequacy of a program ought to be conceived primarily in terms of its effectiveness in attaining defined educational objectives, rather than by specifying the format, sequencing, or other aspects of the means whereby these ends are achieved. For example, the Editors believe that the most justifiable approach at present is the use of relatively small-step programs with an emphasis on the role of emitted responses as instrumental in learning, rather than techniques which use student responses merely as a means of assessing whether learning has taken place from reading or study. But we also believe that it would be very unwise, at the present time, for quality-control standards to reflect this opinion. For example, standards should not at present specify that responses must occur every so often or must be of a particular form, or that a program must have specified feedback characteristics or step sizes. Rather, performance characteristics in terms of the effectiveness with which the student is taught should at present be the basis of any set of standards.

ORGANIZATIONAL AND ECONOMIC CONSIDERATIONS

It seems likely that the design and utilization of classrooms will need to undergo some change as increasing use is made of auto-instructional methods. The consequences of the wider use of these methods for the administration and organization of school systems need to be examined. Only now is there the beginning of an attempt to try out programmed learning materials in the context of a total school system and to study their impact on instruction across the entire range from elementary school to the university level. It is interesting to speculate upon the extent to which experimental teaching machines and programs may need to be adapted to fit into existing educational practices, and the ways in which the latter, conversely, may have to change in order to utilize fully the advantages of these new methods. Some individuals venture to predict that classes as such may disappear, grading systems may change markedly, and even the physical construction of classroom buildings may be altered in the course of this adaptive process.

The manner in which auto-instructional programs are to be developed also has potentially far-reaching implications for educational organization. Will there be as many programs as there are textbooks in a specific subject? In view of the cost of developing programs, it seems very unlikely that this can be afforded. It is important to recognize that teaching-machine programs are, like educational television and films, relatively expensive to produce as compared with textbooks and other traditional learning tools. Also, because of the fact that much of their virtue lies in their perfectability through developmental tryout, and in the feedback this provides to the programmer, their development is a time-consuming as well as an expensive process. It is also a process which calls for marshaling of diverse skills and for special training in concepts and methods of programming. In many cases, a full-scale programming effort is likely to require the coordinated efforts of teachers, students, psychologists, subject-matter specialists, technicians, and educational administrators. For these reasons we can afford to be less profligate in the construction of individual learning programs than in the production of traditional educational materials. The fact that this will mean fewer competing programs for any given subject matter emphasizes the need for better standards of quality control and evaluative criteria.

The diversity of skills and concentration of effort required for the development of programs seems to imply somewhat increased centralization of production. However, it is neither desirable nor inevitable that such centralization reach the stage where a few agencies have a monopoly in the production of programs. Indeed, at the present time the trend is anything but this; currently, a number of programs are already being written at various locations around the country in each of several subject matters, e.g., Spanish, French, Russian, algebra, arithmetic, physics, and spelling. At present, the more difficult aspect of the problem is not so much to avoid a centralized, totalitarian control of educational-materials production as it is to conserve our developmental resources. There is the opportunity for stimulating diversity, but there is also need for coordination so that effort is wisely apportioned.

It is also important to insure that the effort expended results in programs which are, by and large, as good as the current state of the art can permit. This in no way precludes the exercise of ingenuity and the trying out of various approaches, but it does suggest the need for a reasonable degree of coordination among programming groups. Such coordination should prevent unwitting duplication of effort in a small number of subject-matter areas at the expense of leaving other subject-matter areas untouched. There is a need for an efficient medium of information exchange, and for making trial programs in developmental stages readily available to other programming groups for experimental use. Such an exchange function relates both to the interchange of specific program materials and to the communication of programming techniques as these are evolved. It may be noted that a similar program-exchange function was early recognized as desirable in digital-computer programming during the period of its enormous expansion during the 1950's, and that effective and mutually beneficial results for a similar kind of informal, working-level interchange among programmers appear to be developing at the present time in the auto-instructional area. It may become desirable to formalize this exchange function somewhat through several central clearinghouse agencies, instead of leaving it on the present basis of personal correspondence and contact.

In addition to the costs of program development, an important economic consideration is the cost of reproducing lengthy programs for wide-scale use and of purchasing teaching machines in which they can be used. The cost of reproduction and distribution of programs should not be underestimated. For example, it has been estimated that reproduction in programmed-text format for a two-semester course in high-school physics might require several thousand pages of material costing as much as $50 a set to print. Amortizing over a five-year period would result in a printing cost of $5 per semester per student, which would appreciably exceed current textbook budgets. Substantial reduction in this level of cost seems possible in several ways, including new printing techniques, use of microfilm program presentation, and, possibly, new developments in programming methods. In any case, program development costs are likely to remain sufficiently high for some time to make it prohibitively expensive, within existing school budgets, to write programs which only few students will use. However, for certain business, industrial, or military purposes which have high per-man training costs, the writing of programs for relatively small groups of trainees can be economically worthwhile.

Present teaching machines range in cost from several hundred dollars apiece down to around twenty dollars or less. We can expect two trends in the future. One will be the mass production of low-cost devices for presenting well tested programs. For example, a machine costing twenty to fifty dollars would be a very reasonable capital investment if it is assumed that a single machine will be used for four or five high-school courses by a student, and will have a life of at least five years. This would mean that, amortized over its expected amount of use, machine cost should not exceed about a dollar or two per student per course, which would appear reasonable in light of prevalent costs of textbook and other teaching aids.

The other trend that can be anticipated in teaching-machine design is experimentation with electronic or other complex machines for experimental or research purposes. It is possible that devices may eventually be designed which permit a wider range of instructional capabilities with less extensive preprogrammed materials. Also, various technical innovations to reduce the labor and cost of programming may appear. If so, relatively greater costs for electro-mechanical or electronic devices might be practicable in the light of the resulting reduction in program costs. The serious application of technology to the auto-instructional concept is only in its infancy, and it is obviously impossible to foresee all of the possibilities which may become technically and economically feasible during the next few decades. What does seem certain is that changes in current instructional practices, materials, and standards are inevitable and that progress may occur in a variety of possible ways. The history of technological development in other fields would suggest that the machines and programming techniques of today are likely to be superseded, within a relatively few years, by improved devices and methods resulting from continuing research and development.

THE OUTCOMES OF INSTRUCTION

If the cost factors discussed above are thought of as the denominator in an index of the instructional efficiency, the numerator for such an index must be the quality of educational achievement that is attained. If the student learns more, we can perhaps agree to pay more for the result. What kind of things can we expect the student to learn more effectively through the use of auto-instructional methods? Concern is frequently expressed over the extent to which their use may result in stereotyped learning and rote performance rather than real understanding of subject matter. This point has been raised explicitly or implicitly in several of the papers in this volume. The answer would appear to rest upon a behavioral definition of what is meant by "understanding." Programmed learning sequences can be set up to provide a carefully organized progression of instances by which the student is led to develop discriminations and generalizations that enable him to learn and use complicated concepts. The student can be taught these concepts in such a way that he can apply them to a variety of situations, can employ them to solve problems, and can utilize them for creative thinking and invention. It is when a student can do this that we are likely to say that he understands a concept. It seems certain that programmed learning sequences can thus shape the characteristics of behavior which not only permit the student to respond to particular situations in particular ways, but which enrich his understanding and develop his creative capabilities.

For attaining these outcomes efficiently it seems apparent that there needs to be a great deal of guidance provided for the responses made by the student at various stages of learning. The demonstrable advantages of "learning by doing" can best be realized only if the student receives sufficient guidance and feedback so that he practices those responses that will lead him to the desired proficiency. Without adequate guidance, much of his time may be dissipated in irrelevant activity and he may lose interest through frustration and lack of adequate reinforcement. Further, he may be practicing inefficient or incorrect patterns of response which later must only be unlearned and corrected. This is as true in academic subjects as it is in skills like violin playing and tennis.

For learning to be effective a quite close degree of control needs to be exercised during early stages while basic skills and patterns of response are being shaped. However, as fundamental stages are mastered it is equally important for guidance to become more flexible so that the student is given adequate opportunity for freer expression and the exercise of individual talent and creativity. (The basic emphasis in programmed learning on guidance and control over conditions that determine the learner's behavior sometimes leads to a concern that the student may develop an undue degree of dependence on being "spoon fed" by the easy progression of steps in a program.) Contributors to this volume have pointed out that the student must learn to cope with an unprogrammed world in which he will have to learn from newspapers, reference books, and personal experiences. Clearly (any good teaching procedure must "wean" the student from dependence on special helps which he is given during early stages of learning, and a teaching-machine program is no exception.) The later stages of a programmed-learning sequence need to be designed to accomplish this. One of the specific goals of a good program, in fact, should be to prepare the student for subsequent learning. It can do this both by developing habits of careful, sustained attention to textual material and by leading him to master fundamental concepts and transferable skills which are necessary as tools for further independent study, laboratory work, or effective group participation.

THE FUTURE

Teaching machines and programmed learning can have a major impact on education. Their use can effectively and dependably guide the student's learning-by-doing as he proceeds, as rapidly as his abilities permit, through carefully pretested instructional programs. It can thus be made economically feasible to provide every student with many of the benefits of a skilled private tutor, since auto-instructional materials can anticipate and be responsive to his needs for mastering each aspect of a subject matter. Not only do programmed materials themselves thus have the potential for producing much more efficient learning than has hitherto been generally possible, but their wise use should make possible the much more constructive use of the teacher's talents.

The basis for consistent improvement in educational methods is a systematic translation of the techniques and findings of the experimental science of human learning into the practical development of an instructional technology. To achieve the full benefits inherent in this concept, instructional materials and practices must be designed with careful attention to the attainment of explicitly stated, behaviorally-defined educational goals. Programmed learning sequences must be developed through procedures that include systematic tryout and progressive revision based on analysis of student behavior.

The rate at which the methods of programmed learning may influence the practices, staff, and facility requirements of school systems is difficult to predict. However, it seems likely that we can look forward to significant revisions in concepts of classroom practices and teacher functions. We also need to foresee and plan for the time when, because of increased instructional effectiveness, a much more advanced range of educational content can be included in the precollege curriculum.

Appendices

Introduction

Appendix I is an annotated compilation of papers in the field of teaching machines and programmed learning. Appendix II is a consolidated bibliography of all the references cited by authors of papers in the book, together with citations for the papers that are abstracted in Appendix I. (Appendix II is not annotated except for the use of a few designating symbols, identified below.)

Content of Appendix I

The abstracts or other annotated entries in Appendix I are intended to furnish a comprehensive compilation of all known published papers as of June 1960, and of most of the known unpublished papers which deal directly with teaching machines and auto-instructional concepts or programs. Also included are some papers which do not deal directly with these central topics, but which are of sufficient relevance to auto-instructional methods that their abstracting in this context appears desirable. These include, for example, selected papers dealing with the use of prompting and active student response procedures in learning from films, and a few studies from the more general literature of experimental psychology, education, and training. However, the listing of papers in these latter categories is a selective rather than a comprehensive one. A complete integration of all the relevant studies of learning and instruction, and identification of their implications for the design of teaching machines and learning programs, must remain a task for the future.

Publication Status and Accessibility

The decision to include citations in the Appendix for a number of hitherto unpublished papers stems in part from the newness of the field and from some of the peculiarities of the manner in which reports of work have been circulated. With a few exceptions, most of the reports have not been formally published until very recently, but have been circulated, largely in mimeographed form, to interested individuals. However, a number of these unpublished papers have been cited in the published literature. For these reasons it appeared that comprehensive coverage of the *de facto* literature in the field could not be provided if no unpublished papers were listed.

Although a number of unpublished research papers have thus been included in Appendix I, it was decided not to try to list the rapidly increasing number of unpublished auto-instructional *programs*, even though some of these have been cited in publications. Rather, the citation of programs is in the main limited to the as yet quite small number of formally published programs.

The problem of making the Appendix citations a useful guide for the reader has been complicated by several facts. Many of the published as well as unpublished papers cited by authors, especially government and foundation-sponsored reports, are relatively inaccessible. There is also considerable overlap and duplication among papers—e.g., those presented at professional meetings or informally circulated in mimeographed form and subsequently printed, sometimes under different titles. Where limited circulation has been given to an unpublished paper which has subsequently been published in substantially the same form, reference to the unpublished version is frequently omitted.

Particularly in the case of certain classes of government reports the publication status is sometimes ambiguous. The fact that a given paper has been classed as "published" does not necessarily mean that it is at all readily available through usual sources. Some of the presumably "published" material is in mimeographed form, bound or unbound, has had only limited distribution, and may be virtually inaccessible at present except through special sources. On the other hand, some presumably "unpublished" working papers have been rather widely distributed in mimeographed or multilithed form and have been widely cited in the published literature. Standards used by different agencies at different times are not consistent, and no hard and fast categorizations are possible. However, as a helpful guide, the symbol "U" (for "unpublished") has been used following entries in Appendices I and II to distinguish reports which have not been released as published documents, and which in most cases can be obtained only through the authors or other special sources. The symbol "L" (for "limited accessibility") has been similarly used to designate papers which, though in some sense published, are likely to be difficult to obtain except through the authors, their sponsors, institutions, or other special sources. The symbol "Δ" has also been used, often in conjunction with the symbol "U" or "L," to designate certain classes of reports, originating in government agencies and related institutions, for which further explanatory information concerning publication status and accessibility is given below in the section on *Availability of Documents*.

Preparation of Abstracts

The annotations in Appendix I vary greatly in length and comprehensiveness, ranging from a sentence or two to several pages of summarization and excerpting. The Editors have had to make a number of arbitrary decisions concerning the degree of detail in which to treat various papers. The preparation of annotations and abstracts has been governed by several considerations. Only title citations (without abstracts) are given for papers which are included in the volume itself. When papers substantially duplicate each other but have appeared in more than one source (sometimes under different titles), mention is generally made of this fact in order to guide the reader and avoid his searching through several sources only to come up with material he has already obtained. One major guide in determining the degree of detail in which papers are abstracted is the directness of relevance to the topic of teaching machines and programmed learning. Extensive summarization and in some instances quoting of fairly lengthy excerpts has been utilized for papers which, for one reason or another, were not included in the text of the volume, but which make significant points not covered elsewhere.

A number of the abstracts have been prepared by individuals other than the Editors, including Sidney L. Pressey, E. B. Fry *et al.*, and, in some instances, the authors of the abstracted papers. Where individuals other than the Editors have prepared the abstract, the name of the abstracter is given. Abstracts signed "*E. B. Fry et al.*" are from Fry, Bryan and Rigney (1960). Where no name is given, the abstracts were prepared by the Editors.

Arrangement and Form of Citation

The form of giving reference citations in both Appendices I and II has in general followed the publication practices of the National Education Association. To some extent, however, these practices have been modified where it appeared useful to do so for particular classes of papers, utilizing conventions employed in publications of the American Psychological Association.

In both Appendices I and II, publications are arranged in alphabetical order by surname of the author or authors. Within a group of two or more publications for any one author or combination of authors, the publications are arranged chronologically, earlier publication coming first. In cases where a given author (or combination of authors) has more than one publication within a given year, parenthetical letters (a), (b), etc., are used to identify the several publications within that year, which are arranged alphabetically by title. These designating letters in combination with the year of publication have also been used throughout the text to identify the particular paper that is being cited.

A single asterisk (*) preceding an entry in Appendix II means that an abstract, annotation, or excerpt for the paper is given in Appendix I. The double asterisk (**) preceding an entry in Appendix I or II indicates that the paper is included in the text of this volume. The page numbers on which the paper appears in the volume are given in brackets at the end of these citations. Where no source is given for such entries in Appendices I and II other than the date (1960), they are papers that appear for the first time in this volume.

Abbreviations

It may be helpful to the reader to identify some abbreviations used to refer to government agencies and professional organizations which are cited in references. These include *APA* for American Psychological Association, *EPA* for Eastern Psychological Association, *MPA* for Midwestern Psychological Association, *WPA* for Western Psychological Assocation. *NEA* refers, of course, to the National Education Association, and *DAVI* to the Department of Audio-Visual Instruction of NEA, the publishers of this volume. *AERA* refers to the American Educational Research Association, a department of NEA.

The following abbreviations of governmental agencies may be helpful for identifying the origin of reports. Army: *HumRRO* refers to the Human Resources Research Office, a contract research agency for the Department of the Army, operated by George Washington University. Navy: *ONR* refers to the Office of Naval Research, with headquarters in Washington, D. C. *NTDC* refers to the Naval Training Device Center, in Port Washington, Long Island, New York. This agency was formerly called the Naval Special Devices Center, abbreviated *SDC*. (These initials are to be distinguished from the initials SDC currently in use to refer to the System Development Corporation in Santa Monica, California.) Air Force: Current Air Force organizations cited include the Wright Air Development Division (*WADD*) of the Air Research and Development Command (*ARDC*), at Dayton, Ohio—formerly designated the Wright Air Development Center (*WADC*). Others include the Air Force Office of Scientific Research (*AFOSR*, or simply *OSR*), with headquarters in Washington, D. C.; the Air Force Cambridge Research Center (*AFCRC*), recently redesignated as the Air Force Command and Control Division of ARDC; and the Operational Applications Laboratory (*OAL*), a component of *AFCRC*. Former Air Force agencies referred to in citations include the Air Force Personnel and Training Research Center (*AFPTRC*); the Human Resources Research Center (*HRRC*)—a predecessor of AFPTRC; the Human Resources Research Laboratories (*HRRL*); and the Human Factors Operations Research Laboratories (*HFORL*). Nonmilitary agencies: These include the United States Office of Education (*USOE*), and the National Academy of Sciences—National Research Council (*NAS-NRC*). Among nongovernmental agencies, *FAE* refers to the Fund for the Advancement of Education of the Ford Foundation, and *AIR* refers to the American Institute for Research.

Availability of Documents

Because of the inclusion in this bibliography of a number of government reports and other papers often not accessible through ordinary library channels, it seems desirable to give the reader some guidance concerning the availability of these documents. Doctoral dissertations, in addition to their availability on interlibrary loan from issuing institutions, can in many instances be obtained on microfilm from University Microfilms, Ann Arbor, Michigan. Many of the authors of papers cited in this volume are psychologists; their current addresses are generally listed in the Directory of the American Psychological Association, or can be obtained by writing to the APA central office at 1333 Sixteenth Street, N.W., Washington 6, D. C.

Formally published reports by Army, Navy, and Air Force agencies can often be obtained, by institutions having Department of Defense contracts, from the Armed Services Technical Information Agency (ASTIA) Document Service Center, Knott Building, Dayton, Ohio. Reports issued by independent research agencies and foundations can often be obtained from the authors, the issuing institution, or the sponsoring organization. In some cases they are also available from libraries. The rather complex picture concerning availability of certain categories of government reports which have been marked with the symbol "$\wedge$" is discussed below.

Technical Reports of HumRRO are considered formal publications and can be obtained in a number of university libraries. Technical Memoranda, on the other hand, are considered unpublished documents and ordinarily can be obtained only from the author or in some instances on loan from the HumRRO office at George Washington University in Washington, D. C.

Reports of the Office of Naval Research (ONR) are frequently republished in professional journals. The original reports, following initial distribution, are not distributed by ONR, and generally can be obtained only through the authors or issuing institutions. In-service or contractor reports sponsored by NTDC and USN-SDC may be available in some university libraries, and in many instances are also available for purchase from the Office of Technical Services, Department of Commerce, Washington 25, D. C. In the case of contractor reports, copies can also sometimes be obtained from the originating institution.

A somewhat more extended commentary on the availabilty of Air Force reports is made here because of the complexity resulting from organizational changes and the number of relevant reports issued by no longer existing Air Force organizations. This applies particularly to the former Air Force Personnel and Training Research Center (AFPTRC), and to several of its component organizations, including the Armament Systems Personnel Research Laboratory (ASPRL), the Training Aids Research Laboratory (TARL), and the Maintenance Laboratory (ML). Technical Memoranda and unpublished laboratory notes or staff memoranda from the latter three organizations are difficult to obtain. Copies of most of them are believed to be available on interlibrary loan from the Government Documents Division of the University of Colorado Library, in addition to their possible availability from the authors. Files of many of these reports are also maintained at the WADD Aerospace Medical Laboratory (Dayton, Ohio) and Personnel Laboratory (Lackland Air Force Base, Texas).

More widely disseminated reports that were published by AFPTRC and HRRC at Lackland Air Force Base include reports classed as Technical Reports, Technical Notes, Research Bulletins, Research Reports, and Development Reports. A file of these reports is maintained at the library of the Personnel Laboratory, Wright Air Development Division, Box 1557, Lackland Air Force Base, Texas. If other sources of availability fail (see below), a copy of a needed report in these published series can usually be obtained from this source by using standard interlibrary loan procedures.

Provision for wider availability of these categories of reports, published at the headquarters of AFPTRC and predecessor organizations (but not of laboratory-published technical memoranda), was made during the operation of AFPTRC by providing single copies of these reports to a considerable number of university and public libraries throughout the country. (The collections at any one institution may not be complete, however.) The depository libraries include the libraries of the Universities of Alaska, California (Berkeley and Los Angeles), Chicago, Colorado, Denver, Florida, Hawaii, Illinois, Michigan, Minnesota, Nebraska, Pennsylvania, Pittsburgh, North Carolina, Oklahoma, Oregon, Puerto Rico, Southern California, Virginia, Tennessee, Washington (Seattle), and Wisconsin. Other depository libraries at which these reports may be available include the libraries of Brown University, Coburn College, Columbia University, Cornell University, Harvard University, the Harvard Graduate School of Business, Idaho State College, Indiana University, the Iowa State College of Agriculture, the State University of Iowa, Johns Hopkins University, Lehigh University, Louisiana State University, Michigan State College, North Carolina State College, Ohio State University, Princeton University, Purdue University, and Tulane University. These depository libraries also include the Library of Congress, the California State Library at Sacramento, the John Crerar Library in Chicago, the Indiana State Library (Indianapolis), the New York State Library (Albany), the Boston Public Library, the Brooklyn Public Library, the Cincinnati Public Library, the New York Public Library, the Milwaukee Public Library, and the Seattle Public Library.

SUMMARY OF SYMBOLS USED IN APPENDICES

* (Used in Appendix II only.) An annotation of this reference is given in Appendix I.

** The reference is included in the text of this volume; page numbers in brackets refer to this volume.

△ A government or other institutional report, the availability of which is discussed in the Introduction to Appendices I and II in the section "Availability of Documents."

U An unpublished paper.

L A "published" paper which has generally limited accessibility.

These symbols are used beginning on page 582.

APPENDIX I

Adkins, D. C. "Measurement in Relation to the Educational Process." *Educational and Psychological Measurement* 18:221-40; Summer 1958.

This paper analyzes the relationship of measurement to education, considering the nature of measurement and the characteristics of the learning process. Skinner's treatment of learning is the framework employed. Skinner emphasizes the fact that education is intended to alter the relation of classes of stimuli to particular classes of behavior. This requires that behavioral goals must be made concrete in terms of specific curricular content. Since this involves a change in the behavior of the learner, the measurement of this change is an important aspect of the educational process. The process involves the following steps:

(a) Defining behavioral goals or objectives; (b) planning stimuli (that is, curriculum materials) and teaching methods; (c) predicting the degree to which objectives can be attained; (d) applying the teaching methods to the selected curriculum content, i.e., adjusting the goals or the content and methods; (e) testing achievement; and (f) defining new behavioral objectives in the light of how well the old ones were met.

These steps are discussed in terms of an interaction between the learning process and measurement problems. Skinner's notions about behavior are discussed in terms of an analysis of measurement procedures. The author agrees with some of the implications of Skinner's analysis as presented in his *Science and Human Behavior* (1953), and disagrees with others. Emphasis is placed on the development of "a more orderly curriculum and a more highly organized plan for measurement, closely geared together at every stage of the educational process."

Allen, W. H. "Research on Film Use: Student Participation." *Audio-Visual Communication Review* 5:423-50; Spring 1957.

Allen summarizes a number of studies in which active student response is used in conjunction with presentation of content by means of films. Graphical presentation of results is included for some of the studies, including those of Hovland, Lumsdaine, and Sheffield (1949), Michael (1951; see also later report by Michael and Maccoby, 1953), Kurtz and Hovland (1953), Jaspen (1950), Gibson (1947), Ash and Carlton (1951), and others. Results and conclusions of studies as presented by authors are summarized under the headings of Verbalization of Response, Perceptual Motor Responses, Knowledge

---

See p. 580 for key symbols used on following pages.

of Results, Mental Practice, and Note-Taking. The author discusses, as main factors in the effectiveness of active student response in instructional programs, the effects of practice vs. motivation, the nature of the stimulus (including ease-difficulty, rate of presentations, and relation to intelligence), and alternative procedures such as repetitive presentation, which could be used with the same increment of time required for overt-responding procedures.

Amsel, A. "Some Comments on the Manipulation of Error Responses and Reinforcement Schedules in Self-Instructional Training Devices." Unpublished Staff Memorandum, Maintenance Laboratory, AFPTRC. Lowry Air Force Base, Colorado, September 1957. △ U

This is an earlier, unpublished version of the following paper by Amsel (1960).

**Amsel, A. "Error Responses and Reinforcement Schedules in Self-Instructional Devices." (1960) [p. 506-16]

Angell, D., and Lumsdaine, A. A. *Prompted Plus Unprompted Trials Versus Prompted Trials Alone in Paired-Associate Learning.* Pittsburgh: American Institute for Research, 1960 (in press). △ L

This experiment was concerned with the use of a partial degree of prompting as compared with complete prompting. The experiment may be considered from two points of view: First, it offers an experimental clarification of the results obtained in previous experiments by Cook and Kendler (1956), Cook (1958), and Cook and Spitzer (1960). In the second place, it can be viewed as an initial experimental investigation of the efficacy of partial or incomplete prompting, in which the incompleteness of prompting is a matter of the *frequency* with which prompting is provided to the learner, rather than (as in subsequent experiments) being a matter of the *degree* of prompting provided to the learner on any one trial. The basic comparison was between two groups, one of which received prompting on *all* trials in learning a set of paired associates, and one of which practiced responding without prompting on every fourth trial. The paired-associate learning materials used and intra-trial stimulus-response time intervals were substantially the same as those employed by Cook (1958). The results showed that, contrary to the interpretation of previous data by other investigators, learning was significantly more efficient under the experimental condition of *incomplete* prompting (prompting on three-fourths of the trials) than under conditions of complete prompting (prompting on every trial). This result is in accordance with the theoretical predictions outlined in the paper by Lumsdaine (1959b), and is in contradiction to the interpretation derived from previous studies, that maximum prompting represents the optimal condition.

Angell, G. W. "The Effect of Immediate Knowledge of Quiz Results on Final Examination Scores in Freshman Chemistry." *Journal of Educational Research* 42:391-94; January 1949.

The purpose of this experiment was to determine the effect of immediate and delayed knowledge of quiz results on three types of learning in freshman chemistry. The types of learning were knowledge of facts and principles, application of facts and principles in nonquantitative problems, and application of facts and principles in quantitative problems. Students in the experimental group obtained immediate knowledge of quiz results by means of the Angell and Troyer punchboard. Delayed knowledge of results was obtained through the use of IBM answer sheets which were scored and returned to the students at the meeting following the taking of the quiz. Differences between scores on a final examination were statistically significant in favor of the experimental group that used the punchboard and received immediate knowledge of results.

**Angell, G. W., and Troyer, M. E. "A New Self-Scoring Test Device for Improving Instruction." *School and Society* 67:84-85; January 31, 1948. [p. 66-68]

Arnoult, M. D. *A Comparison of Training Methods in the Recognition of Spatial Patterns.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, February 1956. *Research Report* AFPTRC-TN-56-27. △

This study deals with methods for training a critical skill component in the performance of radar bombardiers—the interpretation of complex patterns appearing on a PPI radar scope. The report presents the results of a laboratory study in which several different methods were used to train observers to recognize complex visual patterns. The study is of interest because of a comparison of "constructed" responses composed (drawn) by the learner, as compared experimentally with responses consisting of verbal description of the stimulus. In terms of performance on a criterion test it was shown that the best results were obtained by instructional methods which involved a number of attempts to draw the pattern from memory. Methods involving a verbal analysis of the pattern in terms of interrelationships among pattern elements were significantly less effective.

**Barlow, J. A. *Earlham College Student Self-Instructional Project, First Quarterly Report.* Report submitted to the U. S. Office of Education by Earlham College, Richmond, Ind., 1959. (Mimeo.) [p. 416-21]

Barlow, J. A. "Project Tutor." *Psychological Reports* 6:15-20; 1960. (a)

The article reviews some of the basic notions underlying teaching machines and their implications for education. The second half of the article presents a preliminary set of questions that a programmer might ask himself in the course of constructing a program for use in a particular course.

Barlow, J. A. A review of *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter). *Harvard Educational Review,* 1960 (in press). (b)

This review gives a concise (approximately 1000 words) factual summary of Galanter's book. The individual papers are briefly summarized.

Beck, J. "On Some Methods of Programming." *Automatic Teaching: The State of the Art* (Edited by E. H. Galanter). New York: John Wiley & Sons, 1959. Chap. 3, p. 55-62.

This paper presents some systematic concepts concerning the preparation of material for an exposure device similar to the type originally designed by Skinner. An essential feature of the teaching process is that a student's responses are guided and restricted in such ways as to facilitate the occurrence of the correct response. This feature has been recognized in the approach employed by many acquainted with the general problem of programming. The means of achieving this general objective are, however, unclear, since a student's responses may be restricted and guided in a great number of different ways. These range from all types of hints and prompts—either formal or based on relevant behavioral processes—to simply presenting the response which it is desired a student acquire. These different ways of restricting responses may be conceptually represented as teaching rules and classified according to the manner in which they restrict behavior.

Teaching rules can differ in the degree to which they restrict behavior, whether they involve stimuli in addition to the discriminative stimuli which set the occasion for the response, and whether they are explicit or implicit. Examples of some broad classes of teaching rules involving the presentation of additional stimuli are copying, analogy, and deductive rules. A copying rule restricts the student to copying one of a specifiable number of alternative responses presented. A student's responses are completely restricted if only one, the correct response, is presented. To take a simple illustration: in teaching spelling, the word to be learned can appear spelled correctly, which the student copies. Analogy and deductive rules present additional stimulus material (such as spelling analogues and spelling rules) from which the student can obtain the correct response by means of drawing an analogy or a logical deduction. A different possibility is to present the discriminative stimulus without any additional stimuli. In this case, a student's response is restricted only by implicit rules of thumb he may have developed concerning different properties of the discriminative stimulus—such as the spelling of certain sounds. All types of prompts connected with the discriminative stimulus—as breaking up a word to be spelled into syllables and accentuating the pronunciation of each syllable—may of course be used.

Many other types of teaching rules can be identified, as for example rules using different kinds of controlled association. Different teaching rules not only restrict responses in different ways but establish different mediating responses and associations. These mediating responses and associations may be taken as a basis for distinguishing types of learning. Rote learning can be thought of as the learning of a stimulus-response connection by means of a copying rule. Thus, the teaching rule to be used can depend on what it is desired to teach, whether it is a specific response—teaching a student to mark a statement true or false; a set of facts—teaching a student a set of conditional responses; or problem solving—teaching a student to make related responses in related situations. The paper characterizes the variables involved in the teaching process and explores how different rules might be used.—*J. Beck*

Besnard, G. G.; Briggs, L. J.; Mursch, G. A.; and Walker, E. S. *Development of the Subject-Matter Trainer.* Technical Memorandum, Armament Systems Personnel Research Laboratory, AFPTRC, Lowry Air Force Base, Colo., March 1955. (ASPRL-TM-55-7) △ L

This is a description of a preliminary form of the Subject Matter Trainer, a later version of which is described in subsequent reports by Besnard, Briggs, and Walker (1955) and Briggs (1958).

Besnard, G. G.; Briggs, L. J.; and Walker, E. S. *The Improved Subject-Matter Trainer,* Technical Memorandum, Armament Systems Personnel Research Laboratory, AFPTRC, Lowry Air Force Base, Colo., April 1955. (ASPRL-TM-55-11) △ L

This technical memorandum gives a slightly more detailed description of the Subject Matter Trainer described in the *Psychological Reports* article by Briggs (1958).

**Blyth, J. W. "Teaching Machines and Human Beings." (1960) [p. 401-15]

Briggs, L. J. "Intensive Classes for Superior Students." *Journal of Educational Psychology* 38: 207-15; April 1947.

This is a journal publication of the following paper in the Ohio State University Press series of abstracts of doctoral dissertations. Substantially the same information is presented as in the 1949 abstract, except that somewhat greater emphasis is placed in delineating the major findings, and there is somewhat more condensation of the original data.

Briggs, L. J. "The Development and Appraisal of Special Procedures for Superior Students." *Abstracts of Doctoral Dissertations* 58: 41-49; 1949. (Columbus, The Ohio State University Press.)

The investigation developed and appraised two procedures for facilitating the work of superior college students in a required five-hour undergraduate course in educational psychology, enrolling about 300 students a quarter. The "accelerate seminar" provided, for about thirty students or roughly the upper 10 percent in ability and previous academic record, one informal two-hour meeting a week instead of the usual five one-hour meetings. The two hours, with much use, especially outside of class, of instructional tests with punchboards were spent in informal laboratory exercises, reports, and discussions. All material in the regular sections was covered by these students, plus some 200 pages of extra reading. Appraisal in terms of scores on objective midterms and final examinations in four such seminars in four quarters showed 65 percent of these students obtaining an A or B grade, as compared to 55 percent of a control group paired with them as to score on an entrance test of academic ability, previous scholastic record, age, sex, and grade in the preceding course in general psychology. Only 28 percent of the 801 students in the 27 regular sections made such high grades. Many of the accelerate seminar students were able, because of time saved, to take an extra course or take on part-time work needed for self-support. A follow-up several quarters later showed the seminar students maintaining their academic superiority over their controls.

For very superior and responsible students, a yet more intensive four-week proficiency-examination program involved one two-hour planning meeting a week, the usual reading plus about 200 pages extra, and intensive self-instructional periods using the punchboard, and tests of various types covering the course. Of 24 students in this program, 96 percent made A or B grades on the final examination, as compared to 83 percent of cases paired with them as were the controls for the seminar group—and 28 percent A or B in the regular sections.

It is concluded that self-instructional methods are especially suitable for superior students; using such methods, they do better even than other students of equal presumed potentialities who are given more conventional methods of instruction.—*S. L. Pressey.*

Briggs, L. J. *A Troubleshooting Trainer for the E4 Fire Control System.* Technical Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo. May 1956. (ML-TM-56-16) △ L (a)

This is a preliminary version of the following published report by Briggs.

Briggs, L. J. *A Troubleshooting Trainer for the E-4 Fire Control System.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, July 1956. Development Report, AFPRTC-TN-56-94 (ASTIA Document No. 09887071) △ (b)

This report describes a device designed for affording active practice, with prompt feedback, in learning to diagnose malfunctions in a complex electronic (airborne fire control) system. The training-equipment configuration includes a microfilm projector somewhat similar to the "tutor" device described by Crowder (1960b; see pages 286-98), coupled to a unit containing synthetic components of the equipment system which the student is learning to troubleshoot. This configuration provides students opportunity for rapid, efficient practice in systematic troubleshooting procedures, so that they develop skill in identifying the smallest replaceable unit of the equipment which causes the malfunction. As in some of the applications of the variable-program microfilm device described by Crowder (1959, 1960b, c; see also Lumsdaine, 1959d, pages 5-22), the sequence of information presented to the student by the microfilm depends on the responses he makes. This allows branching sequences in which the results of various checks for a complex equipment-system malfunction may be simulated, as well as permitting contingent sequences to correct errors made by the student. The report gives a fairly complete description of the general configuration and function of the device, illustrated by photographs, but does not present details of mechanical or electrical design.

**Briggs, L. J. "Two Self-Instructional Devices." *Psychological Reports* 4: 671-76; December 1958. [p. 299-304]

A preliminary presentation of this paper was given, under the title "Two More Self-Instructional Devices," at the 1958 annual convention of the American Psychological Association, Washington, D. C., September 1958.

Briggs, L. J. "Applied Research Problems." Paper read at a symposium on automated teaching, Western Psychological Association Convention, San Diego, Calif., April 1959. U (a)

Programming techniques and teaching machines now being developed for use in education are relevant to, but not sufficient for, dealing with many applied industrial and military training situations.

Using the example of training of maintenance personnel for complex electronic equipment, it is shown that three stages of development are needed in planning training for a given job: (a) identification of the verbal learning to be accomplished; (b) programming of this information into teaching machines; and (c) development of complex "simulators" to provide practice in actual job tasks in order that the verbal information is transferred to and becomes "consolidated" with a smoothly functioning perceptual and motor performance of the actual job operations required.

Steps (a) and (c) above are given special attention to emphasize that an analysis of the criterion performance itself does not identify the verbal skills needed, and that our present knowledge of "pure examples" of learning, (such as rote memorization, concept formation, and conditioned responses), does not provide a realistic framework for development of the needed complex skill trainers. Specific research questions are cited to illustrate the nature of the applied research which would contribute to solutions of training problems for electronic technicians.—*L. J. Briggs.*

**Briggs, L. J. "Problems in Simulation and Programming in the Design of Complex High Skill Trainers." Paper read at Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Annual Convention, Cincinnati, September 1959. [p. 329-35]. (b)

Briggs, L. J. "Teaching Machines, Education, and Job Skills." *Psychological Reports* 5: 210; 1959.(c)

This article is a one-page overview of teaching machine developments and applications.

Briggs, L. J. "Teaching Machines for Training of Military Personnel in Maintenance of Electronic Equipment." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons; 1959. Chap. 12, p. 131-46. (d)

Training of personnel in maintenance of electronic equipment appears to require that the trainee learn many kinds of things, some of which do not lend themselves readily to practical classification and description, either in psychological or in subject-matter categories. Even so, in general terms, most such maintenance training must include some learning activity roughly describable by the following phrases: (a) learning of independent paired-associates, such as the name and the appearance of physical objects; (b) learning of fixed sequences of actions to be performed, such as a series of steps in performing inspection or adjustment procedures; and (c) learning to select and use appropriate information, and to make logical deductions from a series of checks in order to solve problems in troubleshooting the equipment.

The Subject Matter Trainer and a card-sort device are described. (For a more detailed description of these devices, see Briggs, 1958.) Two devices described and discussed more fully [see also Briggs and DuVall, 1957] are:

1. A complex procedural trainer, in which the physical aspects of the equipment and its functions are simulated, and in which automatically-programmed sequential instructions are employed along with feedback signals, in order to provide the trainee automatic instruction in fixed sequences of activities. In this simulated maintenance situation, written instructions, such as would be provided in a field manual, are presented as stimulus materials. As the trainee performs, the system "freezes" if an error is made, and the source of the error and the acceptable corrective action are indicated. Upon resumption of correct procedure, the system "unfreezes," and the next instruction is presented. Thus the device affords guided practice in the manual accomplishment of complex, fixed procedures, for which instructions in written form would be available in the field situation. For appropriate sequences, provision is made for memorization of the stimulus materials also, so the trainee can learn to perform without dependence upon the manual.
2. A troubleshooting trainer, having simulation features similar to the procedural trainer, with means for affording the trainee practice in making wise sequences of checks while attempting to isolate a malfunction. This device, by the use of logic circuits, indicates after each check whether it is a reasonable check, considering the total series of preceding checks.

Each of these proposed devices is discussed in terms of its scope of possible use, its known limitations, either in training function or type of equipment, and its potential saving in instructor personnel and trainee efficiency.—*L. J. Briggs.*

Briggs, L. J. *A Survey of Cueing Methods in Education and in Automated Programs.* Pittsburgh: American Institute for Research; Research Report AIR-314-60-IR-106, May 1960. (AFOSR-TN-60-286) △ L (a)

This is the first of a series of reports on cueing methods prepared under a project sponsored by the Air Force Office of Scientific Research. As background for some of the experimental work under the project, it discusses methods used by teachers and writers of autoinstructional programs in cueing or prompting students to maximize the likelihood of their giving the correct responses to instructional exercises. The use of direct prompts versus less direct forms of cueing is discussed and illustrated. The relative absence of evidence on cueing technique from studies in the experimental laboratory is noted. Teacher techniques in starting with minimal cueing and proceeding to more direct prompts are pointed out, and prompting features of some training devices are described. Crowder's intrinsic programming method is discussed from the standpoint of cueing. Specific cueing techniques used in "Skinner-type" programs by Holland and others are illustrated and briefly commented on.

Briggs, L. J. "Teaching Machines." *USAF-NRC Educational and Training Media: A Symposium.* (Edited by G. Finch.) Washington, D. C.:

National Academy of Sciences—National Research Council; 1960 (in press). (b)

This report was originally presented at the USAF-NRC-sponsored Symposium on Training and Education Media, Washington, D. C., August 18 and 19, 1959, under the title, "Teaching Machines: History and Possible Applications to Air Force Education and Training Programs."

It traces the development of teaching machines, beginning with Pressey's early machine and ending with Skinner's disk devices and such "paper-teaching" machines as Crowder's book.

Briggs feels that the use of teaching machines is justified whenever it is shown to be superior to regular training methods. These training devices should also be used when there is a shortage of skilled instructors. Briggs identifies a number of factors which, he feels, makes teaching machines more effective than normal instructor-student interchanges. They are: (a) an expert program writer can reach a number of students, (b) misconceptions held by minimally qualified instructors will not be passed on to students, (c) errors are immediately corrected and do not lead to further errors in the problem sequence, (d) the student works at his own rate, (e) slow students are required to master the material and fast students can save time, (f) slow students are not embarrassed by their lower rate of learning, (g) fast students can do extra assignments, participate in discussion groups, and perform laboratory or field exercises to develop further skills, (h) the need for examination is greatly reduced, (i) the experiencing of success augments motivation, (j) good instructors may use their time to better advantage than performing rote drill.

Briggs states that Pressey's and Skinner's devices present essentially "fixed programs." In the future, he believes, machines will retain some of their present characteristics, but may first run a test program to find out what the student already knows on the topic, and then select one of several basic programs available on the same lesson. The machine may be used to adjust the program on a continuous basis as a function of the responses of the student. This type of "branching" is already used to some extent in Crowder's Scrambled Books, the Tab Test, and the Microfilm Device. In conclusion, Briggs anticipates the possible need for more than one kind of machine because of the probability that there may be more than one kind of learning. Individual differences among students may justify other kinds of flexibility; also, in some subjects, at least, there is a need to wean the student from initial training stages.

A section is devoted to further information on the Subject-Matter Trainer and a report of two informal tryouts of this trainer conducted at Air Force bases, by training supervisory personnel. Lack of suitable measures of performance proficiency led to inconclusive results in comparing experimental and control groups. However, student and instructor responses to the device were favorable. A negative recommendation concerning routine use of the device was based in part on a shortage of personnel to handle maintenance.—*E. B. Fry et al.*

Briggs, L. J., and Besnard, G. G. "Experimental Procedures for Increasing Reinforced Practice in Training Air Force Mechanics for an Electronic System." *Symposium on Air Force Human Engineering, Personnel and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. (Library of Congress Card Catalog No. 56-60063) p. 48-58.

This is a report of an AFPTRC study which made an experimental comparison between two training procedures, in one of which an automated teaching device was used. Two groups of Air Force students matched on aptitude test scores were given a portion of an Air Force maintenance course. The control group was taught by "appropriate" conventional instructor-presented methods. In the experimental group, by contrast, a considerable portion of the training time was devoted to reinforced practice, utilizing the Subject-Matter Trainer described by Briggs, 1958 (see pages 299-304) and other devices and techniques to increase active practice by the students. Analysis of proficiency performance scores by a covariance analysis in which student motivation was controlled showed that the performance of the experimental (reinforced practice) group was significantly higher than that of the control group. However, a significant difference was not found between the two treatments in terms of written tests, even though these correlated positively with the performance test.

Briggs, L. J., and DuVall, W. E. *Design of Two Fire Control System Maintenance Training Devices.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, September 1957. Technical Report AFPTRC-TR-57-7. (ASTIA Document No. 134242). △

This is a description of the rationale and preliminary engineering design of two complex devices for procedural and troubleshooting training; see Briggs, 1959(d).

Briggs, L. J.; Plashinski, D.; and Jones, D. L. "Self-Pacing Versus Automatic Pacing of Practice on the Subject-Matter Trainer." Unpublished Laboratory Note, AFPTRC, ASPRL-LN-55-8, Lowry Air Force Base, Colo., September 1955. (Mimeo.) △ U

This is an unpublished report of an experimental study which compared self-pacing in the presentation of self-instructional materials (subject works at his rate, as in the usual teaching-machine situation), versus controlled (automatic) pacing in which the instructional items or frames were presented at a fixed rate determined in advance by the experimenter. Two groups of subjects of approximately 30 subjects each were given about 13 minutes practice on a 20-item paired-associate learning task with the Subject-Matter Trainer (see Briggs, 1958), employing the "single-error-permitted mode." For the self-paced group, the number of trials (repetitions of the material) varied from one to seven in the time allowed, depending on the speed at which the individual subject worked. The automatically-paced group received three trials, with the paired associates presented at a fixed 13-second interval. The test scores of the automatically-paced group were not significantly different from those of the self-paced group; in fact, the means for the two groups on a test given at the conclusion of training were almost identical ($t = .23$). However, an interesting finding was that the automatically paced group was significantly poorer in terms of accuracy on the first practice trial ($t = 5.4$), though this difference washed out on subsequent trials. The authors' interpretation is that it takes a while for a subject to adjust to the automatic, externally-paced procedure.

Brown, G. I., and Hodgkinson, H. L. "A Note Concerning 'An Application of Recent Developments in Psychology to the Teaching of German,' by C. B. Ferster and S. M. Sapon." *Harvard Educational Review* 28: 156-57; 1958. [This commentary on the paper by Ferster and Sapon is reprinted here with minor editorial changes. —Eds.]

One of the crucial problems in education today involves the efficient utilization of teacher time. Mechanical teaching devices have been suggested, and seem to hold some promise, as a partial solution to this problem. Ferster and Sapon have presented a provocative approach to the more efficient teaching of a modern language through the use of mechanical teaching devices. Their approach is commendable in that it applies psychological theory to a concrete teaching problem. At the same time, there appear to be certain methodological inadequacies in the study.

Ferster and Sapon provide descriptive detail for the original sample of 28 subjects, but no similar descriptive detail for the final sample of six subjects. We know from the study that the original sample contained "4 undergraduates, 5 law students, . . ." etc., but nothing is reported as to who remains in the final sample. Also, no reasons are given for the unusually high (three-fourths of the sample) dropout. The comment given by the authors on this point was that "those students who failed to finish the course stopped at the more difficult lessons."

Ferster and Sapon state that the mean time for teaching German by their method was 47.5 hours, and calculated the mean time for a conventional semester course as 48 hours. Minor criticisms of their comparison of the time spent in the two methods are: (a) time for use of the preliminary tape recording is not included, apparently, in the data of the study, and (b) class hours being 50 minutes in length instead of 60 reduces the average time in the conventional semester class from 48 to 40 actual clock hours.

Perhaps of greater consequence are the following: (a) The 22 dropouts who would probably have increased the mean time of the study were not considered in arriving at the mean time of 47.5 hours, which was based on the final sample of six subjects; (b) There is a good possibility that these six composing the final sample, if given the opportunity to do so in a conventional class, might have mastered an equivalent content in considerably less time than the 40 hours ordinarily spent in a semester course; (c) a possible effect on the authors' comparison of time spent in the two methods exists in the fact that the students of the sample are not typical. They were volunteers, while students in the conventional class situation are often there because it is required. It would seem that the motivation of the volunteer sample would be greater, thus affecting both time spent on, and amount of progress in, the material to be mastered.

The authors provide, in a description of their findings, much detail as to means and ranges. However, no statistical interpretation of the significance of these findings is provided, because of the lack of comparison with any specific comparison groups. The authors are not able to provide this comparison because the findings of the study are not based on standardized tests; they are based on tests ". . . designed to measure mastery of a given corpus . . . felt to have greater specific validity than currently used standardized tests, inasmuch as the latter necessarily attempt to match current textbook and classroom materials and procedures." (p. 65) However, specific validity is of little value as opposed to that obtained from standardized tests, because the study is an attempt to compare the new procedure with the more conventional procedure. In order to compare the two methods, the evaluative instruments must have general validity for the measurement of achievement under both methods.

Another questionable assessment of their results is the statement: "Inasmuch as the material represents a large and adequate sample of the German language, completion of the work sheets is in itself a meaningful measure of accomplishment, since in reaching criteria in each unit the student has in fact written the correct German sentences from the English stimulus material." The assumption is implied that the content is automatically learned and will be retained when completed on a work sheet with only two successful identifications.

It would have been helpful in the interpretation of the findings if the statement, "Those students who failed to finish the course stopped at the more difficult lessons," had included an explanation of what is meant by "more difficult lessons." This is of particular importance in relation to the authors' statement: "A graded level of difficulty was attempted so that the progression from item to item is so slight that the student seldom fails." This obviously was not the case for three-fourths of the original sample; consequently, information as to what constitutes a more difficult lesson should have been reported. It is certainly possible that two identifications are simply not enough for retention.

Perhaps the most disturbing interpretation can be found in the authors' concluding paragraph: ". . . the results show the possibility of teaching basic foreign language skills with high efficiency and economy of time for both student and teacher." In view of the analysis presented in this note, it would seem that the results of the study lend little support to this conclusion. This research is of little value in justifying the use of mechanical teaching devices. Such research hinders rather than furthers the acceptance of the use of mechanical teaching devices in the classroom. It is also unfortunate that the inadequacies of this research will lend ammunition to those who are emotionally indisposed toward the acceptance of devices which in all probability, if judiciously utilized, would be of much assistance to the classroom teacher.

Bryan, G. L.; Bond, N. A., Jr.; LaPorte, H. R., Jr.; and Summers, S. A. *The AUTOMASTS: An Automatically-Recording Test of Electronics Troubleshooting.* Technical Report No. 11, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles. August 1957. (Mimeo.) △ L

The "Automasts" test described in this report is an automatically recording form of the "Masts" tests described by Grings *et al.*, (1953). A fairly detailed description of this device is given by Fattu (1956).

Bryan, G. L., and Rigney, J. W. *An Evaluation of a Method for Shipboard Training in Operations Knowledge.* Technical Report No. 18, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, September 1956. △ L

This report presents the results of a study performed to evaluate a possible shipboard training method. The standard multiple-choice item was modified to convert it to a training device. This was accomplished by developing short verbal explanations to accompany each alternative for each item. These explanations were concealed under paper tabs so constructed that a tab opposite each alternative could be pulled off when that alternative was selected as the correct answer for the item. This method of presentation is sometimes referred to as the "pull tab" device. In this fashion, it was possible to relate the subject matter of the item to the trainee's information, and to explain to him why the alternative he selected was the correct or incorrect answer.

Forty-eight senior NROTC students were randomly assigned to three groups. One group served as a control and answered the items in the usual testing way. During the training session they made exactly one response to each of 150 items. A second group was given immediate knowledge of the correctness of their responses by a "right" or "wrong" and was required to continue selecting answers to each question until they located the keyed answer. The third group was the same as the second with the important exception that explanations of why each item alternative was right or wrong were printed under pull-off tabs opposite the alternatives.

All groups were tested on a multiple-choice criterion test one week after the training session. Comparison of the criterion means indicated that explanation of choices was significantly more effective than the other modification, and that a self-trainer incorporating this training method also could be scored and used as a diagnostic test. —*E. B. Fry et al.*

Bryan, G. L.; Rigney, J. W.; and Van Horn, C. *An Evaluation of Three Types of Information for Supplementing Knowledge of Results in a Training Technique.* Technical Report No. 19, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, April 1957. △ L

This study sought to determine the relative effectiveness of three different kinds of explanations when employed in the multiple-choice trainer format. Report No. 18 by Bryan and Rigney (1956), had indicated that explanations generally were effective in this training situation.

A selection of multiple-choice items was made from among those employed in the original study. Three different kinds of explanations were written to rationalize the scoring key used for each item alternative. The kinds of explanations employed differed as follows:

Type 1. Explanations which gave the trainee the correct definition or description of the chosen alternative.

Type 2. Explanations which indicated the *principal reason* why the chosen alternative was keyed as correct or incorrect.

Type 3. Explanations which pointed out the probable *operational consequences* of the course of action represented by the alternative.

Forty-eight senior NROTC students from Stanford University served as experimental subjects. Each student participated in a one-hour training session which was followed one week later with a criterion test. The primary experimental data consisted of the differences between each man's score before and after receiving the training.

Analysis of the data revealed that significant learning occurred as a result of the one-hour training session with all of the three types of explanations employed. No one type of explanation was found to be superior.—*E. B. Fry et al.*

Bryan, G. L., and Schuster, D. H. *An Experimental Comparison of Troubleshooting Training Techniques.* Technical Report No. 30, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, December 1959. △ L

The training effectiveness of guidance and explanations were investigated using 115 Navy electronics technician students. The students were assigned randomly to five different training groups, defined below. Each group was trained to apply troubleshooting principles in locating defects in a malfunctioning DAS-3 Loran, an electronic navigational aid. In the experiment, the Loran was simulated by means of schematic diagrams, illustrations, and circuit status information.

Each student participated in the experiment for five days. The first was devoted to standardized formal training consisting of tape and slide lectures, question periods, and familiarization with training materials. This was followed by three days of training, and one day for the criterion test.

Group 1 used the same troubleshooting training materials as did the other groups, but were given no assistance. They worked through as many of the 60 training problems as they could in three days. . . . Provision was made to "test" a part if it was suspected of being defective. The problem was concluded when the defective part was located, and the man went on to the next training problem.

Group 2 followed the same procedure as Group 1 but in addition it employed a specially constructed device called the Optimal Sequence Trainer (OST) which provided immediate knowledge of results and delayed cueing. Before making a circuit check the student set a code number into the OST. If that number corresponded to a programmed sequence of troubleshooting steps a green light showed. The student was permitted to look up the results of the check only if the machine had given him a green light. If for any reason the student delayed for more than seven minutes, the OST would tell him which step to take.

Group 3 followed the same procedures as the others except that the guidance was controlled by a Scrambled Clues List. The student worked through the problem by following a list of steps specified by a troubleshooting expert. These were incorporated in a scrambled list so that the trainee had to follow each step in proper sequence.

Group 4 was treated as Group 2 and in addition was provided with written explanations to be opened after each "green-light" check. The explanations sought to justify the selection of that step in the optimal sequence.

Group 5 was similar to Group 3 except that members were given explanations to accompany each of the steps in the optimal sequence.

Results of the criterion test showed that the two guidance groups were significantly superior to the no-guidance group (.05 level) for the total number of items correct and the total number of tabs pulled. The criterion test was a modified Pull-Tab Test (see Bryan and Rigney, 1956). No statistically significant differences among guidance groups occurred in terms of weighted errors or elapsed time.

The groups receiving explanations were significantly superior (.01) to the others in terms of the total number of items correct, number of tabs pulled, and weighted errors. Again, elapsed time differences were insignificant.—*E. B. Fry et al.*

Buitenkand, N. *Methods of Instruction in Component Location for Troubleshooting in Complicated Equipment and Means for Testing Accuracy of Answers.* U. S. Patent No. 2,764,821. Filed August 25, 1953 (assigned to Van Valkenburgh, Nooger, and Nelville, Inc., New York), patented October 2, 1956.

This is the basic U.S. patent for the Trainer-Tester (see Renner, Inc., 1958).

Cantor, J. H., and Brown, J. S. *An Evaluation of the Trainer-Tester and Punchboard Tutor as Electronics Troubleshooting Training Aids.* (George Peabody College), Port Washington, N. Y.: Special Devices Center, Office of Naval Research, October 1956. Technical Report NTDC-1257-2-1. △ L

In this experiment, conducted at Great Lakes Naval Training Station, three groups were used. The first was the T-T, or Trainer-Tester group; the second was the P-T, or Punchboard-Tutor group; the third was the E-group, or Equipment-Only group. The Trainer-Tester was a large sheet with silver overlay printed over answers to troubleshooting problems. (This device was developed by Van Valkenburgh, Nooger and Neville, Inc., its early development having been done by the USN Special Devices Center.) The Punchboard-Tutor was essentially the same device as that developed by Pressey and reported in his 1950 study. The Equipment-Only group used the Navy's traditional method of training in electronics using actual mockups of electronics equipment. All three groups were enrolled in regular equipment troubleshooting courses. In general, the Trainer-Tester and Punchboard-Tutor groups were superior to the Equipment-Only group. The report states that the T-T and P-T groups were definitely superior in intellectual aspects while the E-group was superior in equipment usage. Other criteria which showed the superiority of T-T and P-T groups were examinations, grades in course work, and lab grades. The report concluded that laboratory exercises can be replaced in part by training aids.

Estimated cost figures were given: figured on a per-student basis in a 10-week course, the cost of using a Punchboard-Tutor was 20 cents, the cost of using a Trainer-Tester was less than one dollar, and the cost of using equipment only was $2.10. These estimates ignore developmental costs and are based on replacement or consumption expense.

An interesting sidelight is that exactly the same problems can be programmed on the punchboard as were programmed on the Trainer-Tester; these, of course, were the same problems as those presented on the equipment.—*E. B. Fry et al.*

Carpenter, C. R., and others. *The Classroom Communicator.* (The Pennsylvania State University) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, October 1950. Technical Report SDC 269-7-14. △ L

The Classroom Communicator is a group-type learning machine. An elaborately wired special classroom was devised in which each seat was equipped with a semi-private (hidden) response button. The instructor at the front of the classroom presented on film a multiple-choice question, and each individual responded at his seat by pushing the appropriate answer button. The response was recorded on the instructor's panel. A cumulative record was kept for each student, as well as for the entire class. Immediate knowledge of results was available to the students by showing the answer following each question. Responses could be timed.

This device lacks the characteristic of rate flexibility, having been developed originally as a film rater or film analyzer in part. There are numerous film analyzers which work on the basis of an individual resistance valve at each seat. The student gives information to the instructor or to the machine by turning the dial on the resistance valve. General Electric Company makes perhaps the best known response analyzer.—*E. B. Fry et al.*

Carpenter, F. "How Will Automated Teaching Affect Education?" *The University of Michigan School of Education Bulletin.* October 1959.

This is an introductory article stating principles underlying teaching-machine instruction. (See also the following article.)

Carpenter, F. "The Teaching Machine." *Recent Research and Developments and Their Implications for Teacher Education.* Thirteenth Yearbook. American Association of Colleges for Teacher Education, Chicago, Ill., February 1960.

Dr. Carpenter (of the University of Michigan School of Education) considers the term "self-instructional device" to be more fitting than the designation "teaching machine." The reason is that "teaching machine" suggests something that can replace the teacher entirely. Since nothing of this sort actually exists, it is more accurate to speak of "self-instructional devices" instead of "teaching machines." The definition given includes any mechanical or electronic apparatus which presents learning material in small segments, each segment calling for a response which is followed by some knowledge of results. This excludes some of the revelant testing devices, such as the punchboards, which do not themselves present a step-by-step program. A description of Skinner's teaching-machine principles is given.

What direct evidence is there for adopting teaching machines on a mass scale in our schools? At present the arguments for using machines are largely theoretical, with very little supporting data of a direct nature. The term "direct data" is defined in terms of comparative evidence from experiments using control groups. Too few of such experiments have been reported. The only defensible position at the moment is suspended judgment. Another shortcoming of reported studies is inadequate sampling. There is a need for well-controlled, large-scale studies in typical school systems. Despite the lack of sufficient data, however, the machine-learning movement appears to be the most promising development in education since the turn of the century.

What are some of the important issues and problems in the area? The controversy between multiple-choice responding and constructed response is reported to be a lively issue. A list of problems for research is given: What is the effect of making errors on human learning? To what extent should a response be overt? How essential is immediate feedback in human learning? Is motivation under machine learning largely due to the novelty effect or is it more permanent? What is meant by reinforcement in machine learning? How much variety in rearrangements should be available to maintain learning? What are the fundamental principles of programming?

Educational implications of theory and research involving self-instructional devices discussed are that: (a) Self-instructional devices promise to relieve the teacher of much time in presenting course content. (b) The teacher will devote more time to the management of media that promote understanding and application of knowledge. (c) The machine record of pupil performance will supply a valuable aid to the teacher for identifying specific learning difficulties. (d) Methods of evaluating student achievement are likely to be modified radically. Marking practices will be revised, since even slow learners can master content by going at their own pace. (e) The teacher will be able to emphasize more individual counseling and guidance. (f) The management of space in school buildings will be modified. Allotted space will increase for private study; space for large lecture halls will probably decrease. This is contrary to implications seen by some for the use of television in education. (g) Secondary teachers will need greater concentration in the subjects they teach, as well as knowledge of related disciplines. (h) Research on self-instructional equipment will interest more academic psychologists in education.

The remainder of the article presents a discussion of assumptions underlying the implications of the Skinnerian point of view, and an analysis of the lecture method as contrasted with the potential of self-instructional devices.

Carr, W. J. "Bibliography on Teaching Machines." Unpublished report. Lewisburg, Pa.: Bucknell University, 1958. U

This bibliography has been incorporated into the following paper by Carr, and the references it included are also listed in Appendix II of this volume.

**Carr, W. J. *Self-Instructional Devices: A Review of Current Concepts.* Wright-Patterson Air Force Base, Ohio: Aero-Space Medical Laboratory, Wright Air Development Center, August 1959. WADC Technical Report 59-503. [Article appears in this volume under title "A Functional Analysis of Self-Instructional Devices," p. 540-62.]

Cook, D. A. "On Vanishing Stimuli in Instructional Material." *Journal of the Experimental Analysis of Behavior* 3: 1960 (in press)

[This short article by D. A. Cook is printed here in its entirety—Eds.]

In preparing instructional material for a programmed curriculum, one of the general problems to be faced is that of changing the stimulus control of a response. A response may first be made through immediate control of a text, as in copying. In subsequent steps of the program, the response should come under the control of other stimuli. Complete removal of textual control in one step may remove all, or too much, control, so that the correct response does not occur. One solution is to provide intraverbal prompts if these have been established, especially where these are relevant to the subject matter. If intraverbal prompts are inadequate or lacking, the technique of "vanishing" the stimulus may be used. In a spelling example discussed by Skinner (1958c), the word *manufacture* is presented and copied, then presented with *a*'s omitted, to be filled in, then with *u*'s omitted. Accompanying prompts say, "The same letter goes in both spaces."

But in many cases, such an accompanying prompt may not be available, and yet it may still be important to fragment the pattern to build up an "internal" chain in steps (which may be quite arbitrary with respect to the units available in the repertoire at that time).

In such cases, another kind of vanishing or "fading" has been tried, which consists in presenting a fragment at successively weaker intensities of physical stimulation until it disappears. It is difficult and costly to do this, especially with complex visual material, and its effectiveness is open to question. Are we to assume that presenting a stimulus nearer and nearer threshold value acts to strengthen the response it controls? An alternative explanation appears equally plausible. For some subjects, this procedure may act, or come to act, as a *warning* that this portion of the stimulus will disappear "soon." In the presence of such a warning, rehearsal to strengthen that fragment will be reinforced. If this is correct, it would be simpler to *teach directly* in the program a warning signal which would signalize the impending disappearance. For visual material, the section might be indicated by an arrow with a circle at its tail end. If desired, the circle could contain a number telling the number of steps in which the fragment will vanish. Such a convention, once introduced, could be applied repeatedly to different kinds of material, and would be easy to prepare. It could also be used in conjunction with autoclitic prompts given *in advance of* the disappearance, such as: "These two letters are the same" (in the case of *manufacture*, for example), or "These are all plus signs" (in a formula), or "The outline of Venezuela" (on a map).

In teaching young children, putting the material into another color may be used to signalize impending disappearance. With auditory materials, a red light might flash coterminously with the critical section. If the method is effective, many varieties of application will suggest themselves to the programmer.

Cook, J. O. "Processes Underlying the Learning of a Single Paired-Associate Item." Unpublished Staff Memorandum, Maintenance Laboratory, AFPTRC, Chanute Air Force Base Field Extension, September 1956. △ U

This is a more detailed report on the study subsequently reported in Cook's article by the same title (1958).

Cook, J. O. "Supplementary Report: Processes Underlying Learning A Single Paired-Associate Item." *Journal of Experimental Psychology* 56: 455; November 1958.

[This short article by Cook is quoted here in its entirety as it appeared in the *Journal of Experimental Psychology*, except for deletion of the tabular data and minor corresponding editorial changes. It is based on work done under ARDC Project No. 7709, Task No. 37305, in support of the research and development program of the Air Force Personnel and Training Research Center, Lackland Air Force Base, Texas.]

As a test of the theoretical model of paired-associate learning previously outlined (Cook and Kendler, 1956), a second paired-associate experiment was run on 160 airmen, with eight groups of 10 in each treatment. As in the previous experiment, the same experimental materials were used, and the same two learning methods: Prompting (stimulus term, then response term, then overt practice of the response term) and Confirmation (stimulus term, overt practice of the response term, and then the response term). The superiority of the Prompting method, and hence of the model that predicts its superiority, had been questioned on the grounds that the highest level of learning achieved in the previous experiment was only 54.8 percent of the material. Had training been continued, it is alleged, Confirmation might have proved superior in the later stages.

To meet this objection subjects in the present experiment were run to mastery with a test after every three of the 36 training trials. In terms of the number of correct responses, as the theoretical model predicted, Prompting proved significantly superior throughout the learning curve. . . . Prompting also proved superior in terms of the number of different "legitimate" responses (a "legitimate" response being any one of the 10 response terms, whether elicited by its proper stimulus term or by one of the others). In the previous experiment this finding appeared only as a tendency but did not achieve significance. Prompting produced more substitution errors early in learning because the subjects knew a larger number of response terms ("legitimate" responses), but the number of substitution errors, as the model predicted, dropped off rapidly. In the Confirmation condition, substitution errors reached their peak later and showed a significant tendency to persist, due to the fact that the Confirmation method makes it difficult for the implicit response to the response term to become attached to the stimulus consequences of the implicit response to the stimulus term.

Cook, J. O. "Research in Audio-Visual Communication." *Research, Principles, and Practices in Visual Communication.* (Edited by J. Ball and F. C. Byrnes.) East Lansing, Mich.: National Project in Agricultural Communications, Michigan State University, 1960. p. 91-106.

This is an overview summary of a number of experimental studies on factors influencing the learning from sequenced instruction, primarily presented by motion pictures or film strips. Some of the studies relevant for the design of teaching machines and programs include experiments on prompting and the role of overt responses by the learner during the course of instruction.

Cook, J. O., and Kendler, T. S. "A Theoretical Model To Explain Some Paired-Associate Learning Data." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication No. 455, 1956. (Library of Congress Card Catalog No. 56-60063.) p. 90-98.

This study and subsequent studies reported by Cook are concerned with a comparison of two procedures, for paired-associate learning, which have considerable relevance to the rationale underlying programming practices. The two procedures are called Prompting and Confirmation. As used by the authors, Prompting means presenting the response term (which the student is to learn to say in response to the stimulus term) with the response-term presentation timed so that it occurs shortly after the presentation of the stimulus term but *before* the learner is required to make an overt response. Confirmation, on the other hand, means that, following the presentation of the stimulus term, the student first tries to give the corresponding response term *without* any prompting; only after he has attempted to do so is the correct response term presented. The Confirmation procedure thus provides either verification or correction of the response that the student has given.

In this experiment a *test trial* (in which the stimulus term only was presented) was given after each three pairings of the stimulus and response terms, so that cumulative learning curves could be plotted. The analysis made by the experimenter showed that the Prompting procedure produced significantly more correct pairings than the Confirmation procedure. To account for the results obtained, the authors propose a theory which asserts that the learning of individual paired-associate items is, in general, a two-stage process: First the learner learns the unfamiliar response terms as such, and then he learns to associate each of them with the corresponding stimulus term. The authors hypothesize that, "This process is mediated by the stimulus properties of the implicit responses the subject makes to the two items constituting a paired-associate item."

In interpreting these findings, which conclude in favor of prompted responses and hence a higher proportion of correct responses during a training program, it is of interest to note that the response used in the experiments was one which was fairly difficult to learn (a nonsense drawing which had to be imitated accurately) rather than one (such as learning English equivalents of foreign words) where the response term is already familiar but merely has to be hooked to the proper stimulus equivalent. Another significant aspect of the procedures used by these experimenters, in this and subsequent experiments, illustrates the difficulties of experimentation in which the process of measurement may alter the learning conditions which it is designed to assess. The use of interspersed test trials after every three prompted trials meant that the Prompting procedure was actually a mixture of prompted and unprompted trials (the latter used for measurement purposes, but obviously influencing the total learning program also). Further experimentation is needed to clarify this factor as it operates in relation to the results of this study and of subsequent studies (Cook, 1956, 1958; Cook and Spitzer, 1960) [cf. Angell and Lumsdaine, 1960]. It is also of interest to compare Cook's results with those obtained by Irion and Briggs (1957), who also concluded in favor of a procedure which gave more control of errors through an increased degree of prompting, and with similar conclusions by Kimble and Wulff (1953).

Cook, J. O., and Spitzer, M. E. "Supplementary Report: Prompting Versus Confirmation in Paired Associate Learning." *Journal of Experimental Psychology* 59:275-76; 1960.

[This short article by Cook and Spitzer is printed here in its entirety as it appeared in the *Journal of Experimental Psychology*, except for deletion of the tabular data and other minor editorial changes. This research was supported by a grant from the Faculty Research and Professional Development Fund of North Carolina State College.]

Cook and Kendler (1956) and Cook (1958) have shown that Prompting (stimulus term, then response term, then overt practice of the response term) is superior to Confirmation (stimulus term, overt practice of the response term, and then the response term) in learning. The present experiment was undertaken to discover some reasons for this superiority. On the basis of a theoretical model (Cook & Kendler, 1956) plus some conditioning principles, it was hypothesized that the superiority of Prompting was due, at least in part, to a shorter between-term delay, the *omission* of overt practice of the response, or both. Since the temporal interval between the presentations of the two terms of a pair is shorter in Prompting (.25 sec.) than in Confirmation (3.25 sec.), and since the longer the S-R delay is, the poorer the learning will be, it was predicted that a short delay would yield a better learning performance than a long delay. Moreover, on the basis of the theoretical model, it was believed that the occurrence of overt practice during the delay interval would interfere with the establishment of the essential S-R connection that must take place over the span of the interval, and might also interfere with the learning of the response term as such.

Using the same materials as in the previous experiments, the hypotheses were tested by means of a 2 × 2 design consisting of the following four conditions (A) Prompting—No Overt Practice, (B) Confirmation—No Overt Practice, (C) Prompting—Overt Practice, and (D) Confirmation—Overt Practice. . . . There were 35 male undergraduates per condition. The subjects were run to mastery with a test after every three of the 27 training trials.

. . . The results, in terms of the number of correct items (the C score), bear out the predictions that Cond. A will be best (because of a short S-R delay and no intervening practice to interfere) and that Cond. D will be worst (because of a long S-R delay and because of the interfering effect of overt practice during the S-R delay interval). The relative positions of Cond. B and C were not predicted because the relative importance of the two variables was not known.

The two variables also proved to have a deleterious effect upon the number of different legitimate responses—excluding duplicates—that a subject produces on a test (*D* score). A legitimate response is any one of the response terms used in the experiment, whether it is elicited by its proper stimulus or by some other one.

The conclusions are that: (a) Overt practice interferes both with learning the response term as such and with connecting it with its proper stimulus; (b) S-R delay interferes with the process of connecting a response to its proper stimulus, but it has no very consistently reliable effect upon learning the response term as such.

Cornell, F. G.; Damrin, D. E.; Saupe, J. L.; and Crowder, N. A. *Proficiency of Q-24 Radar Mechanics: III. The Tab Test—A Group Test of Troubleshooting Proficiency.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, November 1954. Research Bulletin AFPTRC-TR-54-52. △

This article is a further description of the "Tab Test" [see Glaser, Damrin, and Gardner, pages 275-85]. It introduces one innovation, which is a simple but effective technique for preserving a record of the sequence of responses made by the trainee. This is accomplished simply by having each tab, as it is pulled, placed on a spindle in sequential order. The article also includes some additional discussion of scoring procedures for items employing the tab feedback procedure.

Corrigan, R. E. "Automated Teaching Methods." *Automated Teaching Bulletin* 1: 23-30; December 1959.

This is one of the numerous brief articles which offers a partial restatement of the general notions of self-instruction by devices and programs. (See Rheem Manufacturing Company, *Automated Teaching Bulletin.*)

Coulson, J. E. *An Experimental Teaching Machine for Research at SDC.* Technical Memorandum, System Development Corporation, Santa Monica, Calif., October 1959. U (a)

This is an amplification of the July 1959 Field Note by Coulson and Silberman. Coulson feels that simple teaching machines such as Skinner's and Pressey's are inferior because they do not change the program based on students' responses, explain errors, nor give more items if the first response is incorrect. He feels these faults can be remedied by the use of a small computer, as is being planned. —*E. B. Fry et al.*

Coulson, J. E. "Programming for Individual Differences." Paper read at Symposium on Automated Teaching, Western Psychological Association Convention, San Diego, Calif., April 1959. U (b)

It has been frequently observed that an important characteristic of teaching machines is their ability to give the student prompt reinforcement for certain responses, thereby causing a modification of the student's subsequent behavior. We refer to this modification of the student's behavior as "learning." Less obvious is the fact that a successful teaching machine is also capable of modifying its own behavior in response to information received from the student. In this way the machine adapts its behavior to meet the unique learning needs of the individual student. This adaptive behavior exists to a limited degree in present teaching machines as illustrated by Crowder's scrambled books, Solartron's key punch trainer, and recent studies performed at IBM using a digital computer. The System Development Corporation is currently conducting a study to determine possible effects of one type of conditional sequencing.

The trend toward increasing teaching-machine adaptability is explored from the standpoint of potential future developments in this direction. Possible features of a "self-programming" teaching machine are discussed, and conclusions drawn as to the machine characteristics necessary to produce these features. Studies should be made of the extent of adaptability desirable in teaching machines. In the application of teaching machines to classroom or industrial situations, gain in teaching effectiveness caused by increased machine flexibility must be weighed against corresponding effects on machine complexity, reliability, and expense.—*J. E. Coulson.*

Coulson, J. E., and Silberman, H. F. "Proposal for Extension of Automated Teaching Projects." Unpublished Field Note, System Development Corporation, Santa Monica, Calif., July 1959. U (a)

The view is held that present designs for teaching machines are limited by their lack of adaptability to individual student needs. Future goals of automated teaching projects would include the development of a computer-controlled experimental teaching machine incorporating a large degree of adaptability. After initial experimentation, a smaller machine which incorporates the essential characteristics of the larger machine would be developed; finally, possible usages throughout all educational systems (including academic, industrial, and military applications) would be investigated.

The authors believe that two years will be required to develop an effective computer-controlled teaching machine and instructional items designed to teach symbolic logic. At the same time, some investigation will be made of the teaching of computer programming. A third year will be used to investigate other subject areas, and to determine general principles of teaching-machine design and item construction. A study of the interaction of teaching machines with other components of the educational system might require another two years, making a total of five years. During the final two years, it might be possible to develop a more simplified teaching-machine design which would still incorporate all essential elements.

The authors estimate that the needed capacity of the computer would be 2000 words of internal storage, and 5000 words of external storage. The speed of computer operation is not critical. Most small computers possess more than enough computational speed to keep up with human responses. What computer is used will depend upon a number of considerations. A very large computer (IBM 709) may be utilized because it is available for general use. Or, one of the smaller computers, such as the Librascope LGP 30 or the Bendix G 15 may be purchased for this specific project. The availability of input-output devices of a suitable nature also is a consideration. —*E. B. Fry et al.*

**Coulson, J. E., and Silberman, H. F. *Results of Initial Experiment in Automated Teaching.* System Development Corporation, Santa Monica, Calif., July 1959. [See p. 452-68; cf. also Silberman and Coulson 1960.] L (b)

Coulson, J. E., and Silberman, H. F. "Effects of Three Variables in a Teaching Machine." *Journal of Educational Psychology* 51: 1960 (in press). This is a journal publication of the 1959 study by Coulson and Silberman which appears on pages 452-68 of this volume (1959b).

Cross, K. P., and Gaier, E. L. "Technique in Problem-Solving as a Predictor of Educational Achievement." *Journal of Educational Psychology* 46:194-206; April 1955.

The authors used a Tab Item Test to study preferences for fact versus preference for principles in what they call the BPT (Basic Problem Test). They found that mathematics students prefer principles. Problems which used "facts" were rated easier, but students choosing these problems made more mistakes in arriving at solutions. The authors concluded that BPT is a better predictor of mathematics grades than are language or nonlanguage IQ. The authors make some interesting comments on the use of the Tab Format:

"The Tab Format of the BPT provides a situation in which the subject is able to select the kinds and amount of information he believes will best enable him to solve problems. This technique might extend to those applied situations where the effects of knowledge of results and subsequent test behavior can be measured as well as a large number of selected personality variables (for example, motivation, rigidity, perseveration). The effects of the ambiguity offered by the more abstract principles might possibly be extended to measure personality."

The Tab Method might well be applied to situations where the initial selection and use of a specific series of procedures is an important aspect of the individual's subsequent training. This study was done in cooperation with the Air Force Personnel and Training Research Center.—*E. B. Fry et al.*

Crowder, N. A. "The Concept of Automatic Tutoring." Unpublished Staff Memorandum, Armament Systems Personnel Research Laboratory, AFPTRC, Lowry Air Force Base, Colo., October 1955. △ U

This unpublished paper, cited by Darby (1959) and by Fry, Bryan, and Rigney (1960), gives a preliminary formulation of concepts which are more fully covered in later papers by Crowder (1959, 1960b, c).

Crowder, N. A. *A Part-Task Trainer for Troubleshooting.* Lackland Air Force Base, Tex.: Air Force Personnel Training and Research Center, June 1957. Development Report AFPTRC-TN-57-71. (ASTIA Document No. 131423). △

This report describes a device developed in the context of research and development efforts to provide an improved basis for the troubleshooting of highly complex electronic equipment by Air Force technicians. It helps to "define the field" of the technician's job in that it provides a comprehensive tabulation of the extremely numerous functional interrelations upon which successful diagnosis of malfunctions can depend. Because of the display and feedback characteristics of this simple device, it can also be used as a practice device which has self-instructional properties.

Basically, the device is a display mechanism which provides a complete and systematic tabulation interrelating a large number of test points with the kinds of indications that would be obtained, at each test point, under a considerable number of operational modes or signal injection procedures. For the particular bomb-navigation system for which the prototype was developed this comprises, for normal operation of the system, a table of 28 columns and 478 rows (28 conditions and 478 test points). The output indications for each cell in this huge matrix were first identified for normal operation of the system. A series of similar tables was then developed for each of a number of possible malfunction conditions. Perhaps the principal significance of this device lies less in its properties as a training device than as an illustration of the possibility of analyzing highly complex interrelationships into a finite, though large, number of specific identified contingencies, which can then be utilized as a basis for developing systematic program sequences to teach the needed behaviors in a complex intellectual skill such as electronic troubleshooting.

Crowder, N. A. "Intrinsically Programmed Materials for Teaching Complex Skills and Concepts." Paper read at Symposium on Automation of Teaching, American Psychological Association Convention, Washington, D. C., August 1958. U (c)

This was a 15-minute oral presentation, similar to later papers by Crowder (1959, 1960 b, c), and can be considered to be superseded by these publications. [See the paper by Crowder (1960b) in the present volume, p. 286-98.]

Crowder, N. A. "Automatic Tutoring by Means of Intrinsic Programming." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: Wiley & Sons, 1959. Chap. 10, p. 109-16.

Crowder is interested in the question, "How can you use error response?" . . . This, he feels, can be done by means of his scrambled book. He states that teaching is a "feedback controlled process." Since individual students vary, their needs will be different; hence, their tutoring should be different.

Crowder believes that multiple-choice methods are easiest to work with. Free response methods are more difficult to handle by automated tutoring processes. He defines "automatic tutoring" as individual instructorless tutoring, presenting material in small logical units, each unit being tested immediately. The test results should then determine automatically subsequent presentation of data, either new or review. . . . One of Crowder's hypotheses is that the teaching of meaningful material requires considerable covert activity on the part of the student, that is, "thinking," and that the form of the response that the student makes to the teaching device is essentially irrelevant.—*E. B. Fry et al.*

[See Crowder's later paper (1960b) on p. 286-98 of this volume for a fuller statement of his views—Eds.]

Crowder, N. A. *The Arithmetic of Computers.* Published by Western Design, a Division of U. S. Industries, Inc., Santa Barbara Airport, Goleta, Calif., 1960. (a)

This is a published program (2 volumes, 480 pages), using the method of "intrinsic programming" in the form of a "tutortext" as described by Crowder (1960b). It deals with elementary number-system concepts and is designed to provide a background for digital computer programmers. An earlier version of the first portion of this program was issued under the title "An Automatic Tutoring Book on Number Systems," referenced by Fry, Bryan, and Rigney (1960).

**Crowder, N. A. "Automatic Tutoring by Intrinsic Programming." (1960). [p. 286-98] (b)

Crowder, N. A. "Automatic Tutoring by Means of Intrinsic Programming." *Proceedings of the Educational Testing Service Invitational Conference*, October 1959, on "The Impact of Testing on the Educational Process." Princeton, N. J.: Educational Testing Service, 1960 (in press). (c)

The essential content of this paper is covered by Crowder's paper (1960b) in the present volume [p. 286-98].

Damrin, D. E., and Saupe, J. L. *Proficiency of Q-24 Radar Mechanics: IV. An Analysis of Checking Responses in Troubleshooting on Tab Test Problems.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, November 1954. Research Bulletin AFPTRC-TR-54-53. △

In using the Tab Item Test the authors formulated seven hypotheses about troubleshooting procedures. They attempted to define procedures used by efficient mechanics as distinguished from those used by less efficient mechanics. This report has implications for organization of the content of troubleshooting training.—*E. B. Fry et al.*

Darby, C. L. *An Annotated Bibliography on the Automation of Instruction.* Research Memorandum, U. S. Army Air Defense Human Research Unit, Fort Bliss, Tex., July 1959. (Mimeo.) △ U

This, one of the first annotated bibliographies, summarized a considerable proportion of the teaching-machine literature available up to the time of its preparation. However, it can now be considered as superseded, in the main, by the more readily available annotated bibliography of Fry, Bryan, and Rigney (1960). Annotations are included both for papers dealing directly with teaching machines and for related papers on studies of instructional films and more general studies of training devices. The Fry, Bryan, and Rigney bibliography incorporates some of the abstracts that appeared in the Darby bibliography.

Day, J. H. "Teaching Machines." Unpublished paper. Athens: The Ohio State University, 1959. U

In this as yet unpublished paper, Day describes the teaching machines and programs of Skinner and Crowder, and discusses the application of programming to the teaching of chemistry. Included in the paper are sequences illustrative of two programs which Professor Day has completed, following the Crowder scrambled-book pattern, for the teaching of basic chemistry. The programs seen by the Editors deal with the kinetic theory of gases.

Diederich, P. B. "Self-Correcting Homework in English." *Proceedings of the Educational Testing Service Invitational Conference,* October 1959, on "The Impact of Testing on the Educational Process." Princeton, N. J.: Educational Testing Service, 1960 (in press).

This paper describes a series of self-correcting homework exercises which are to be used in teaching high-school English courses. A plan is proposed by which these self-correction exercises can be used to teach some of the technical aspects of English, and thus cut down the number of class hours required. This should leave the pupil more time for library reading and group discussion.

Dowell, E. C. *An Evaluation of Trainer-Testers.* Headquarters, Technical Training Air Force, Keesler Air Force Base, Miss., June 1955. (Report No. 55-28) △ L

This is a large scale evaluation of the Trainer-Tester. Some 26,000 Trainer-Testers using super-heterodyne receiver circuit as the curriculum were used in training at several Air Force bases. The experimental population consisted of four matched groups of students in the 13th week of basic electronics training. The groups were identified as follows: (a) The control group, which received instruction in troubleshooting on the actual super-heterodyne receiver only; (b) Experimental Group 1, which received instructions in troubleshooting and troubleshooting problems on both Trainer-Testers and actual equipment; (c) Experimental Group 2, which received instructions in troubleshooting and practiced troubleshooting on problems on the Trainer-Tester only; and (d) The pretest group, which received general instructions in troubleshooting but did not practice troubleshooting by any method.

The criterion measure consisted of three repersentative troubleshooting problems on actual receiver equipment. Evaluation of results was based on (a) whether or not the defective part was identified and (b) the amount of time required. In addition, a questionnaire was employed to obtain student reaction. Dowell found that Experimental Group 1, which used both equipment and Trainer-Testers, received the highest average scores on the criterion measure. He also found that the achievements of the control group were significantly better than the pretest group; however, they were not as high as the Experimental Group 1. Experimental Group 2, which used only Trainer-Testers, did not achieve significantly better results than the pretest group that had no opportunity to practice trouble shooting. . . .—*E. B. Fry et al.*

Edward, W. D. "Skinner's Teaching Machines." Unpublished Laboratory Note, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., May 1956. (Mimeo.) △ U (a)

This paper is based on an interview by Edwards with B. F. Skinner in the Spring of 1956. It was rather widely circulated and frequently cited, prior to the availability of published reports of Skinner's work at the Harvard Laboratory after his initial 1954 paper. In addition to describing the several kinds of machines with which Skinner and his associates had been working, Edwards gives several comments which are of sufficient interest to be excerpted here.

Concerning the principles of machine design used by Skinner Edwards says: "The advantages of using machines for Socratic-type teaching, which of course could also be done by a teacher or a book, are that they provide immediate feedback (information about whether the student is right or wrong), that they permit the student to work at his own rate and thus are adaptable to wide ranges of intelligence, that they are self-motivating, and that they offer convenient methods for providing reinforcement. Skinner reports that all students on whom he has so far tried the machines find operating them so much fun that no further reinforcement is needed, and he is not now doing any work on the question of how to use reinforcements in connection with such machines. From his extensive work on reinforcement schedules, however, he feels that the most advantageous reinforcement schedule for such a purpose would be a variable-ratio schedule (i.e., the student is given a nickel or some other extrinsic reward every nth trial, with n varying around some mean value). The typical slot-machine pattern, in which there are fairly frequent small rewards and a very infrequent big reward seems especially well suited to maintaining a high rate of machine operation.

"The most vigorous restriction Skinner feels should be imposed on the design of teaching machines is that the student should compose his answer, rather than recognizing it. This means that multiple-choice questions are inappropriate—a fact which greatly complicates the problem of designing appropriate machines. The reason for this is that reproduction rather than recognition is what you wish to train for, and he does not believe that recognition training transfers to reproduction very well. A self-scoring machine [cf. Skinner, 1958c] solves the problem of accommodating compose-the-answer questions on complicated subjects. Unfortunately, it violates yet another principle, namely that the machine shouldn't reveal the right answer in the course of telling the student that his answer is wrong."

In relation to programming Edwards says: "Of course the most important question about teaching machines is: How should they be programmed? Skinner feels that he doesn't have any very satisfactory answers to this question; this is where research is needed. But some principles do seem clear. Teaching machines, Skinner feels, should teach by a Socratic method of questioning, with a minimum of exposition and a maximum of active participation by the student. The step between each question or item of information and the next should be very small, so that the student is seldom wrong. This is important both to maintain motivation and to make sure that the student gets a complete understanding of the material being taught. The material should be presented in a logical progression so far as possible, and in such a way that the student understands what he is learning, rather than learning by rote. (These are Skinner's words, not mine. If you ask him what 'understanding' means, he will probably say that it means having a very rich associative context for each concept taught before the next one comes along. The implications of this principle for programming teaching machines are obvious, even though the concept of 'understanding' is one of the least well understood concepts in psychology.)

"If you ask Skinner how to produce a program, he would say: Decide what behavior you wish to produce in the student, analyze it into responses, and then teach those responses one by one. It is from this fundamental principle that Skinner deduces the principles which say that the material should be presented in a logical step-by-step progression and that successive steps should be very small.

"In the course of deciding on a succession of steps, Skinner uses another principle. Elements should not be taught before operations; rather operations should be taught as early as possible. In arithmetic, for example, Skinner starts with the numbers 0 and 1. Then he teaches 2, and immediately teaches that $1+1$ is 2, and $2-1$ is 1. Then he teaches 3 and 4, and immediately starts teaching multiplication and division. Then he teaches 5, and starts teaching the concept of prime numbers. In short, he teaches quite complicated mathematical concepts about numbers before the student knows what all the numbers are. This is a technique for producing the 'rich associative context' I mentioned above.

"In the course of trying to ensure that the student is very seldom wrong, Skinner encourages lavish initial use of crutches (my term, to which Skinner has some objections). For example, he may set out the answer in a distinguishing color in the early stages of learning something, or he may write the question as a rhymed couplet with the right answer being the rhyming word in the second line. He feels that such artificial aids, which of course must eventually drop out, are essential in the early stages of learning about something. The principle underlying this is that any road sign, however artificial, is helpful in a trackless wilderness, but becomes unnecessary when the wilderness becomes familiar."

Edwards' own reactions at the time include the following comments:

"I believe that the general notion of using teaching machines is the most important development in applied psychology since aptitude tests—indeed it is potentially the most important development since intelligence tests. Most of the principles Skinner presents seem very reasonable to me. In particular the notion of composing the answer rather than recognizing it seems to me important enough to justify the inconveniences in machine design which it imposes. . . . Another fact about Skinner's work on teaching machines should be recorded, though it is hard to make a general principle out of it. Skinner uses a very great deal of ingenuity, a kind of ingenuity well worth copying, in making up programs. . . . A tentative standard for good programs is that they should interest experts in their field by their imaginativeness without causing the students they are intended for to make too many wrong answers."

In discussing the significance of Skinner's work for Air Force technical training, Edwards comments as follows: "Since Skinner has been mostly concerned with training grade-school students, some modification of his principles to apply to older students and more difficult material is probably called for. But some of these principles still make sense, notably those of small steps, extensive use of crutches in early stages (if this actually is helpful), low probability of error, and operations mingled with elements. Skinner feels that even in electronic training composing rather than recognizing the answer is crucial; to me this seems to be an important and straightforward experimental problem.

"The greatest need in the further development of teaching-machine techniques is, of course, the development of principles of programming. Skinner is surely right in saying that there is no substitute in programming for digging into the actual subject matter to be taught; 'You'll never teach electronics until you have analyzed what electronic behavior is.' But there must be many more as yet unknown general principles of optional programming; research to discover them is badly needed. Another kind of needed work is development of a compose-the-answer machine which can interpret answers, rather than checking them symbol by symbol. The problem is essentially the same as that of developing a translating machine to translate one language into another, and is probably at least as difficult as the translating-machine problem."

Edwards, W. D. "The Use of Statistical Decision Functions in Making Practical Decisions." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. (Library of Congress Card Catalogue No. 56-60063) p. 115-24. (b)

This paper discusses a recurrent methodological problem encountered in experiments for comparing different instructional methods or media—namely, the selection of an appropriate statistical rationale underlying the basic logic of the comparisons. The ordinarily used null-hypothesis-rejection model does not, of course, logically permit any outcome except suspended judgment when negative results are obtained. Thus, the *superiority* of one form of programming over another, for example, can be established—at least within the limits of the particular situation studied—but the *equivalence* of the two cannot be formally demonstrated. Use of fiducial limits to permit a conclusion that the difference, if any, resulting from the two methods is no greater than some specified amount can, of course, be resorted to. A more thoroughgoing approach is to formulate the question as to comparative effectiveness in terms of a decision-theory model which (if reasonable estimates of the necessary parameters can be made) can lead directly to a conclusion concerning the method of choice for specified values of the parameters—even though an equally sensitive null-hypothesis test might have led to failure to establish a significant difference. Edwards discusses some aspects of rationale for application of the decision-theory approach to experimentation on instructional devices, and presents hypothetical examples which are worked out in such a way as to identify what requirements would have to be met to make utilization of this approach possible in a practical experimental situation.

Eigen, L. D. *The Construction of Frames of an Automated Teaching Program.* Collegiate School Automated Teaching Project, New York, N. Y.: November 1959. (Mimeo.) U

Single frames of a program have three components: (a) the *operation* requiring the student to do something, (b) the *data* information he will use, and (c) the *instruction* which is usually "go on to next frame" but may be "see teacher" or "see book." Eigen also classifies frames into (a) *panels* which simply supply information, (b) *review*—no data, just an operation and instruction, (c) *instruction* only, and (d) *forced frame*—student is forced to respond correctly by obvious nature of answer. Cues suggested by the author include color, spatial relations, emphasis (underlining), and demonstrating (arrows and other devices).—*E. B. Fry et al.*

Eigen, L. D., and Komoski, P. K. "Research Summary No. 1 of the Collegiate School Automatic Teaching Project." Report to the Fund for the Advancement of Education, June 1960. (Mimeo.) U

The research described in this report was designed to investigate two aspects of automated teaching: (a) The relative effectiveness of mechanical teaching machines and programmed textbooks as teaching devices; and (b) the effect of grade level on learning by means of these automated teaching methods. A total of 74 male subjects from the Collegiate School in New York City (25 ninth-grade students, 23 tenth-grade students, and 26 eleventh-grade students) were used. The machines used were prototypes of a Skinner-type write-in machine. The programmed textbooks were mimeographed and contained seven items to the page, with questions and answers on separate pages requiring two page turns per frame. The program used was *Sets, Relations, and Functions*, by Lewis D. Eigen. This program consists of 707 frames.

On the day following completion of the program, each student took two tests: a post-test on the material covered in the program, and a transfer test comprised of items which were not specifically covered but which demanded the concepts and skills developed by the program for successful completion. An additional measure available for each subject was the amount of time the subject took to complete the program.

Statistical analysis of the results failed to show a reliable difference in the amount of learning produced by presenting a teaching sequence by a programmed text or by a mechanical write-in machine. This was true for both the post-test and the transfer test. The means of the machine group were lower than the means of the programmed text group. The mean time for the machine group exceeded the time for the programmed text group by more than 37 minutes. However, this time difference is attributed to mechanical difficulty in the machine, which is being improved in production materials. The authors indicate that although the initial cost of a mechanical device might be considerable, it would probably be less expensive to use mechanical devices than programmed texts over a long period of time. A program in a machine is subject to less hard wear by students than the constantly handled text. However, the programmed text has several advantages if it is used for purposes of research when planning new programs and for adapting automated teaching to individual schools.

Within the limits of this experiment, grades 9-11, no significant effect of grade level upon learning was found. This is interpreted to mean that the contention that certain materials are best taught at a given grade level needs serious re-examination. However, the results indicated that transfer is significantly affected by grade level. This latter result is explained by the fact that the eleventh-grade students, having studied more mathematics, have more background on which to base transfer than do the earlier grades. There seems to be little effect of grade level on the length of time needed to complete a program. The authors caution that, although the conclusions as stated in this paper have referred in general terms to automated teaching, they have generalized from results for a program in teaching mathematics.

Emeson, D. L., and Wulff, J. J. "The Relationship Between 'What Is Learned' and 'How It's Taught.' " Technical Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., December 1957. (ML-TM-57-32) △ L

This paper reports several small experiments conducted to explore implication of a rationale which the authors state as follows: "Whenever the learning task is a difficult one, and formalized training materials are justified, there are at least two objectives to be achieved in developing training materials. One of these is to insure active practice; the other is to set up training conditions that will foster appropriate or useful practice. Both of these objectives must be achieved if the training procedure is to be effective. . . . It is not difficult to insure active practice when one designs a formal training program—but it is difficult to insure *appropriate* practice. That is to say, when one restricts the procedures a student may use in learning to those which are part of a formalized method devised to insure practice, the student may be deprived of opportunity to use idiosyncratic methods and past learning that will be useful for the individual learner. This, in turn, may result in a loss of training effectiveness. . . ."

For their experiments Wulff and Emeson used a relatively difficult paired-associate learning task, which required students to learn the names of eight electrical circuits. One of the difficult aspects of the task was that the eight circuits consisted of four pairs, in which the members of each pair were almost identical and thus highly confusable. In an initial experiment, unstructured self-study was compared with a formalized paired-associate learning procedure in which the circuit diagrams were presented one at a time and the learner selected the correct name from a list of all eight names (with order of presentation and arrangement of names in the list both changed from trial to trial, and with an immediate correction procedure used on each trial.) Both self-study and this formalized paired-associate training were relatively inefficient, with the difficulty of learning under these circumstances indicated by the fact that a typical subject was still responding incorrectly about a fourth of the time after nearly an hour of training. Moreover, the learning was considerably more rapid under conditions of unstructured self-study than it was for the paired-associate method, with 11 out of 14 students achieving a perfect score at the end of 40 minutes of self-study. These results, like those of Newman (1957) emphasize the necessity for analyzing a learning task and insuring that appropriate practice and cueing arrangements are provided in a formalized learning sequence if an attempt to develop and apply a science of teaching is to surpass the performance attained through the individual student's own art of learning.

Several subsequent experiments are also reported by Emeson and Wulff, which attempt to implement critical factors in the requirements for effective learning so as to improve on the basic paired-associate technique employed in the first experiment. Because of the confusability of the pairs of circuits, the authors inferred that "it would be necessary for a trainee to evolve a reliable [differentiating] cue for each circuit before he could consistently give the correct name response," whereas if he were "required to attempt to practice name responses before he had established the reliable cue for each circuit he would be highly likely to learn interfering responses." A second critical requirement identified was that of providing practice in responding to reliable cues. In one of the experiments to implement these requirements, a card-sorting procedure was used, prior to associative training, in which the students sorted the stimulus terms (circuit diagrams) into boxes and concurrently learned a number corresponding to the position of each sorting box. The paired-associate training which followed utilized the number cues learned during the sorting procedure. This resulted in considerably better performance, in terms of time to learn, than the simple paired-associate procedure.

In another experiment, two paired-associate procedures were compared which differed only in that for one experimental group the stimulus terms were presented *in a fixed order*, to provide associative support through position cues during the early stages of practice. These students performed better in later stages of practice in which the standard paired-associate arrangement was used (with scrambled order) than did students who had used the standard scrambled-order arrangement throughout the entire period of practice.

Evans, J. L. *An Investigation of "Teaching Machine" Variables Using Learning Programs in Symbolic Logic.* Doctor's thesis. University of Pittsburgh, 1960. (Typewritten.) [Also in press (mimeo.), to be issued as a report under a grant to the University of Pittsburgh from the Cooperative Research Program of the U.S. Office of Education.] U

Programs were designed to teach the construction of short deductive proofs involving 15 rules similar to those used in symbolic logic. Several experimental versions of this two to three hour program were developed in order to investigate certain programming variables. The program variations were the following: (a) A program constructed by the "ruleg" system (see Homme and Glaser, 1959a and Evans, Homme, and Glaser, 1960), compared with a program constructed in a less systematic fashion; (b) a program in which the learner was permitted to use a review card as he went through the program as compared with one in which no review panel was available; (c) different modes of responding (overt vs. covert responding, immediate vs. less-immediate feedback, multiple-choice responding vs. constructed responding). Criterion tests given to the students to evaluate the effect of the program variations were a true-false test, a recall test involving the different logical rules, and a test requiring the student to give actual short deductive proofs. Dependent measures were the time spent on the learning programs, the time spent on the performance tests, and the number of errors made on the performance tests.

The following conclusions were drawn: (a) Experimental variations in mode of responding significantly affected learning time. Subjects not required to make an overt response to each item completed the learning program in about 65 percent of the time required for composed or multiple-choice responding. (b) Criterion performance in terms of error scores was not significantly affected by mode of responding and was not significantly affected by differences between overt and covert responding. (c) A program constructed by the "ruleg" procedure produced, in less learning time, criterion performance comparable to that obtained with the less systematic program. (d) Subjects who did not respond overtly to learning programs took significantly more time on performance tests which immediately followed the program than did subjects who made their responses overtly. Such differences, however, disappeared in a retention test after one week. (e) Performance on a retest given a week later suggested that retention differed as a function of the performance measured. Error scores on the true-false test decreased significantly. Error scores on the recall test showed slight but significant increases. On the tests involving the actual criterion behavior of deductive proof, no significant changes were found between immediate tests and retention tests. An outcome of this research project was the development of a standard learning program for experimentation. The program was designed to be as free as possible from previous mathematical experience.

**Evans, J. L.; Glaser, R.; and Homme, L. E. "A Preliminary Investigation of Variation in the Properties of Verbal Learning Sequences of the 'Teaching Machine' Type." Paper read at the Eastern Psychological Association Annual Convention, Atlantic City, N. J., April 1959. [p. 446-51]

Evans, J. L.; Homme, L E.; and Glaser, R. *The Ruleg (Rule-Example) System for the Construction of Learning Programs.* A report prepared under the Cooperative Research Program of the U. S. Office of Education at the Department of Psychology, University of Pittsburgh. 1960. (In press, mimeo.)

This report elaborates on the ruleg system presented in the article by Homme and Glaser (1959a) on page 486-96. After a preliminary description of the characteristics of the ruleg system, the article proceeds with a discussion of each of the following steps in the construction of a program: (a) Specify the criterion behavior. (b) Write down all the ru's (rules) that it will be necessary to teach the student. (c) Collect all necessary stimulus support in the form of text, notes, advice from subject-matter experts, and so on. (d) Make a preliminary ordering of the way in which the ru's will be presented in the program. The conventional order may not be the best order in programmed learning. (e) Make a ru matrix. This is a listing of the ru's on both a horizontal axis and a vertical axis. The cells where the different ru's intersect indicate to the programmer the possible interverbal connections that may be useful in teaching the subject matter and in providing prompts for learning. (f) Consider each cell in the ru matrix, and construct examples for the program. These examples, or eg's, as they are called, must be chosen on the basis of two primary considerations: first, an adequate number of them should be employed in order to give the student enough practice; and second, a full spectrum of eg's should be used so that the student can consider a wide variety of instances of the rule being developed, e.g., special cases, limiting cases and trivial cases. (g) Number the cells in the original ru matrix to indicate the order in which the ru's are to be presented and exemplified. At least use this as a first approximation. (h) Beginning with this, now proceed to assemble the ru's and eg's into program frames. To help in accomplishing this, frame types are listed, which consist of various types of ru-eg combinations and combinations of incomplete eg's and incomplete ru's which the student must complete or to which he must respond. For example, it is pointed out that ru + eg

$+ \widetilde{\text{eg}}$ (eg tilde, an incomplete example) is the frame type that may provide the best method for the economical introduction of a new rule. It is also pointed out that analogy frames are useful; such frames consist of eg $+ \widetilde{\text{eg}}$. Induction frames are discussed in which, from the first eg, the subject must induce the ru and then apply it to an incomplete eg ($\widetilde{\text{eg}}$). Technical vocabulary words and terms are typically introduced by a ru $+ \widetilde{\text{ru}}$ frame in which the ru is presented and then the student must use the technical word in an incomplete rule. Two other frame types discussed are (a) the $\widetilde{\widetilde{\text{eg}}}$ (eg double tilde), which is essentially a very incomplete example in which the student performs in the presence of minimal cues, and (b) the $\overline{\text{eg}}$ (eg bar), in which the student must respond to a false example and recognize it as a false example. This might be called error-detecting behavior. (i) Using the ru matrix as the basic plan, assemble the frames of various types into a program. (j) Give the program to students and "item analyze" their responses to each frame. (k) Revise the program on the basis of student's responses and comments. In this process of program revision, one first danger to be avoided is over-prompting. Also, it is important to insure that prompts are faded, so that toward the end of a program the student responds to minimal cues, and thus actually performs the criterion behavior desired. (l) Repeat the administration of the program to the students, and repeat the revision until the program is producing the desired criterion behavior reliably and efficiently.

Evans, J. L.; Homme, L. E.; Glaser, R.; and Stelter, C. *Descriptive Statistics: A Programmed Textbook.* (Vol. I of a Program in Statistics.) Albuquerque, N. M.: Teaching Machines Inc., 1960.

This "programmed textbook" [see Glaser, Homme and Evans article on p. 437-45] is the first formally published Skinner-type program, so far as is known to the Editors. It comprises approximately 800 frames, and covers descriptive statistics as follows: The Basic Concepts of Statistics — 78 frames; Frequency Distributions — 84 frames; Measures of Central Tendency—90 frames; Measures of Dispersion—97 frames; Introduction to Normal Curve, Correlation and Variance—142 frames; Variance II—59 frames; Variance III—99 frames; Normal Curve II and Z-Scores—72 frames; Correlation II—57 frames; Correlation III—45 frames. The program materials were developed in the course of detailed tryout with high-school and college students.

Fattu, N. A. *A Catalog of Troubleshooting Tests.* Research Report No. 1, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, by the Institute of Educational Research, Indiana University, Bloomington, December 1956. △ L

This useful 130-page catalog surveys tests in the following groups: (a) on-the-job measures, (b) performance tests, (c) simulator tests, (d) ARC (automatic recording of checks) or tab-type tests, and (e) paper-and-pencil and oral interview tests. Most of the items described were developed under military contract. The catalog mentions Grings' Masts Test and the Automasts Test developed by Bryan and others. Also included is the Automatic Rater device (punchboard), the Trainer-Tester, the troubleshooting board developed by Glaser and Phillips at the American Institute for Research, as well as other studies by Briggs and Besnard, and Cantor and Brown, which are mentioned in this bibliography.

Fattu mentions numerous sources of interest to compilers of this type of bibliography. Chief sources are: (a) the Armed Services Technical Information Agency Title Announcement Bulletins, (b) Human Resources Research Office Bibliography, (c) Psychological Abstracts, (d) Review of Educational Research, (e) information obtained from the Personnel and Training Branch of the ONR, (f) Special Devices Center of the ONR, (g) Human Resources Research Center, (h) American Society for Engineering Education, (i) Technical Education News, and (j) Industrial Arts and Vocational Education.—*E. B. Fry et al.*

Fattu, N. A. "Training Devices." *Encyclopedia of Educational Research.* Third edition. New York: The Macmillan Co., 1960.

This article devotes considerable attention to teaching machines, and describes a functional schema somewhat similar to that presented by Rothkopf (1957). In addition, it briefly indicates some of the characteristics of military training devices which are mentioned in this Appendix, including a brief mention of the "automasts," the "MAC Trainer," and related devices. It also contains a more general discussion of training devices. Fattu discusses these devices in relation to procedures and techniques for educational research.

**Ferster, C. B., and Sapon, S. M. "An Application of Recent Developments in Psychology to the Teaching of German." *Harvard Educational Review* 28:58-69; 1958. [p. 173-85]

Finch, G. (editor). *Educational and Training Media: A Symposium.* Washington, D. C.: National Academy of Sciences—National Research Council, 1960 (in press).

This volume contains several papers of interest to those interested in research on teaching machines, particularly the 1956 paper by Briggs, abstracted elsewhere in this Appendix. This symposium volume publishes papers which were originally presented at the USAF-NRC-sponsored symposium on Training and Education Media, in Washington, D. C., on August 18-19, 1959.

Finch, G., and Cameron, F. (editors). *Symposium on Air Force Human Engineering, Personnel, and Training Research.* Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. Library of Congress Card Catalog No. 56-60063. (Also identified as ARDC Technical Report 56-8, by the Air Research and Development Command.)

This, the first of several related volumes, was prepared on the basis of papers presented at an Air Force sponsored symposium at the National Academy of Sciences—National Research Council, November 14-15, 1955. It contains several papers dealing with factors in the design of training devices. Several of the papers from the 1956 NAS-NRC volume which are especially relevant to research on teaching machines and programmed learning are separately abstracted in this Appendix. These include the 1956 papers by Briggs and Besnard, Cook and Kendler, French, and Edwards.

Finch, G., and Cameron, F. (editors). *Symposium on Air Force Human Engineering, Personnel, and Training Research.* Washington, D. C.: National Academy of Sciences—National Research Council, Publication 516, 1958. Library of Congress Card Catalog No. UG633S96.

This second NAS-NRC volume (see above entry) is based on a symposium held in Washington, November 15-16, 1956. This volume also contains several papers of general interest in relation to the design of instructional material and training equipment, including the paper by Maccoby and Sheffield which is abstracted elsewhere in this Appendix.

Finn, J. D. "Automation and Education: I. General Aspects." *Audio-Visual Communication Review* 5:343-60; Winter 1957. (a)

This paper, the first of a series of three (see below) is summarized by the author as follows: "An attempt has been made in this paper to describe automation, both as an industrial process involving flow and control and as an intellectual system involving planning and decision-making. General effects on education in the areas of curriculum and program of studies, administration, general education, organization for instruction, the necessity for teaching new patterns of thinking, vocational and technical education were discussed. The concept of the possibility of automatizing the instructional process itself was introduced, and the outlines of a new content for education as a discipline were suggested."

Finn, J. D. "Automation and Education: II. Automatizing the Classroom—Background of the Effort." *Audio-Visual Communication Review* 5:451-67; Spring 1957. (b)

Dr. Finn has summarized this second paper of his three-paper series as follows: "This paper has attempted to deal with two aspects of the problem of automatizing the classroom through audio-visual means. These aspects were: (a) the institutional background in which the earlier foundation support of the audio-visual movement by the Payne Fund and the Rockefeller Foundation was compared to the present, much more extensive support of the mass audio-visual approach to teaching supported by the Ford Foundation, its several Funds and the organizations which it underwrites; and (b) the theoretical position requiring audio-visual automation which was shown to be based on the teacher and plant shortage facing this country in the next two decades of swelling school and college enrollments . . . necessitating technological solutions to problems of class size and building space."

**Finn, J. D. "Automation and Education: III. Technology and the Instructional Process." *Audio-Visual Communication Review* 8: 5-26; Winter 1960.

[An abridged version of this article appears in this volume under the title "Technology and the Instructional Process," p. 382-94.]

Freeman, J. T. "The Effects of Reinforced Practice on Conventional Multiple-Choice Tests" *Automated Teaching Bulletin* 1:19-20; September 1959. (See Rheem Califone Corporation, *Automated Teaching Bulletin.*)

Freeman used an electric version of the Pressey punchboard. When the subject inserted a penlight stylus into the correct hole, the light on the pen lit up, indicating that the response was correct. The total number of responses made to get a correct one could be tabulated.

In a large introductory psychology course the punchboard was used by an experimental group taking two 50-item exams, approximately four weeks apart. Their performance was compared on the final exam with a matched control group.

An analysis of items on the final exam which were relevant to earlier specific reinforced exams revealed that there were no effects of reinforcement. Freeman's . . . conclusion is that, under typical conditions of classroom testing, the casual use of reinforced practice has little effect upon performance.

In another experiment performed with advanced psychology students previously reinforced in a similar type situation, items previously reinforced were compared with nonreinforced items. The differences, although not statistically significant, favored the reinforced items. Freeman suggested the possibility that students of lower ability level were benefited more by the reinforced practice.

In a test-retest situation where the experimental group received reinforcement and the control group did not, the mean gain for the reinforced group was five points and for the control group, .15—a ratio of approximately 30.

In another part of the experiment, identical items were used in similar testing, except that one-half of the test was composed of objective items and the other half of completion items. Freeman found that reinforcement procedures resulted in significantly better performance on the objective items but no significant difference on the completion items.

In another section of the study, Freeman attempted to investigate the differences in "systematic" programming versus programming without any system. He feels he has demonstrated that there are some benefits to be derived from a systematic program.—*E. B. Fry et al.* (See Rheem-Califone, *Automated Teaching Bulletin.*)

French, R. S. "Evaluation of a K-System Troubleshooting Trainer." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. (Library of Congress Card Catalog No. 56-60063.) p. 160-65.

This report gives a brief description of the "MAC-I Trainer," which was designed at the AFPTRC Maintenance Laboratory as a device to synthesize functional characteristics of a highly complex electronic system, in such a way as to provide practice in troubleshooting procedures. ("MAC Trainer" stands for Malfunction and Circuitry Trainer.) Although not a fully self-instructional device (much of its use requiring supervision by an instructor), it is closely related to some of the self-instructional devices for electronic troubleshooting described elsewhere, and was designed to provide prompt feedback or knowledge of results to the trainee as he made simulated checks. An evaluation study is also described, in which performance for groups who used the device as part of their training was compared with that of a comparison group which did not use the device.

French, R. S., and Martin, L. B. *A Flight-Line Troubleshooting Trainer for a Complex Electronic System: The "MAC-II Trainer."* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, July 1957. Development Report AFPTRC-TN-57-106. (ASTIA Document No. 134227.) △

This report describes a revised version of the "MAC Trainer" (see preceding report by French, 1956), in which more realistic simulation of components is employed and other improvements are introduced. The report also contains references to a series of more detailed technical memoranda on related devices.

Fry, E. B. "Teaching Machine Dichotomy: Skinner Versus Pressey." Paper read at the American Psychological Association Annual Convention, Cincinnati, September 1959. (In press, *Psychological Reports.*) (a)

Contrasts are shown in several areas affecting teaching machines. (a) Skinner prefers constructed responses while Pressey prefers multiple-choice responses. (b) Skinner prefers that the student make little or no error while responding; Pressey permits some error. (c) Pressey is more interested in having his devices supplement regular instruction, while Skinner and his followers in some instances advocate that teaching machines replace traditional instruction methods. (d) Skinner feels that bright and dull students progress equally well and are treated the same in terms of teaching-machine programs, while Pressey would have completely separate programs for very superior and superior students.—*E. B. Fry et al.*

**Fry, E. B. "Teaching Machines: An Investigation of Constructed Versus Multiple-Choice Methods of Response." *Automated Teaching Bulletin,* December 1959, p. 11-12. (b)

This is a brief abstract of Fry's dissertation study, a more detailed version of which is presented on pages 469-74 of this volume. (See Rheem Califone Corporation, *Automated Teaching Bulletin.*)

Fry, E. B. "Teaching Machines: The Coming Automation." *The Phi Delta Kappan* 41: 128-31; October 1959. (c)

This article is an introduction to the field of teaching machines for teachers. It contrasts teaching machines with audio-visual devices, and points out such interesting characteristics as flexible rate of presentation, immediate knowledge of results, and the fact that teaching machines are in harmony with progressive educational practice which uses only positive rewards and emphasizes activity. . . . Machines are seen as a valuable aid in educational research, since they can eliminate the "teacher variable" when comparisons are to be made of different methods. Of special benefit is the use of machines in the instruction of exceptional children (both bright and dull) as well as in small schools whose curriculum offerings are limited.—*E. B. Fry et al.*

**Fry, E. B. "A Study of Teaching Machine Response Modes." (1960). [p. 469-74.] (a)

Fry, E. B. *Teaching Machines: An Investigation of Constructed Versus Multiple-Choice Response Modes.* Doctor's Thesis. Los Angeles, University of Southern California, 1960. (b)

The essentials of methods used and results obtained by Fry in his dissertation study are included in the published version of the paper which appears on pages 469-74 of this volume.

Fry, E. B.; Bryan, G. L.; and Rigney, J. W. "Teaching Machines:

An Annotated Bibliography." *Audio-Visual Communication Review* 8:1-80; Supplement 1, 1960. 80 p.

This 80-page publication consists of four parts. The first two sections are short, and summarize the growth of interest in the teaching-machine area as represented by reported studies during the past decade. A tabulation of a number of teaching-machine programs completed or under development is given, as well as a tabular summary of experimental studies in this area. Various teaching machines and related devices, both commercially available and others known at the time of writing, are briefly described. The main part of the report (Part III: pages 18-74) is an annotated bibliography of literature in the field. Many of the papers abstracted by Fry are contained in the present volume. All of the other relevant published papers are abstracted in the present appendix, which in some instances utilizes abstracts prepared by Fry *et al.* and in other instances substitutes abstracts prepared by the Editors.

Gagné, R. M. "Teaching Machines and Transfer of Training." Paper Read at Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Convention, Cincinnati, September 1959.

This was one of several papers presented at a symposium at the 1959 meeting of the American Psychological Association. Gagné approached the question of research issues in the design of automated teaching devices from the standpoint of transfer of training. He questioned whether the Skinnerian concept of response shaping by instrumental conditioning is fully adequate to furnish a model for conceptual learning with teaching-machine programs. He summarized his views as follows:

"I have tried to show why I think the conception of 'shaping behavior' by means of teaching-machine programs is a minimally useful one. It appears to me to denote no more about the conditions for concept learning than that when one measures learning one is dealing with behavioral change. I should be inclined to seek a bridge between conceptual learning and the body of knowledge about human verbal learning, including the methodology and data of transfer of training. The latter approach should make it possible ultimately to account for all kinds of human learning within a single theoretical framework. In addition, it is able to encompass the phenomena of behavior generalizability and set as events of importance to concept learning and concept use. From this point of view, the procedure of introducing "small steps" in a learning program is seen as a matter of stimulus and response variation to enhance generalizability. The question of "understanding" of learned material becomes a question, not of predictability of response, but of breadth of utilization of concepts acquired in a learning program, to be determined by observing transfer of training to tasks requiring the use of these concepts."

Gagné, R. M., and Bolles, R. C. "A Review of Factors in Learning Efficiency." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 2, p. 13-54. [This paper was earlier given limited distribution as a separate report prepared for the Air Force Office of Scientific Research, November 1958; AFOSR-TN-58-924; ASTIA Document No. AD 162275.]

As indicated by the authors, "The main body of this report deals with principles derived from laboratory findings and theoretical writings that seem relevant to the problem of promoting efficient learning. In describing these factors and their origin, we have not attempted to provide a detailed review containing all relevant references. Instead, reference is usually made to a key paper, set of studies, or to an authoritative work summarizing the field. Consideration is given to the kind of evidence which exists, the suggestions it has for learning efficiency, and its adequacy in dealing with this problem. Following this review, some conclusions are drawn regarding the present state of knowledge of these principles and the research questions which appear to be in need of investigation."

The main criterion of learning efficiency is defined as the fact that it leads to a high level of performance in a transfer situation. The authors differentiate between several kinds of tasks (especially those of interest to the Air Force), and state that they will pay closest attention to sources of evidence pertaining to the performance of identification, procedural, and concept-using tasks. They also refer occasionally to studies of motor skills.

Manipulatable conditions of learning are discussed under two main headings, readiness factors and associative factors. Readiness factors include motivation (including level of aspiration), reinforcement, set, and attention (intent to learn). Associative factors include the nature of associations (including stimulus similarity, response similarity, mediating stimuli, meaningful similarity, and similarity in serial tasks); intratrial factors (including number of stimuli, number of responses, and meaningfulness); and intertrial factors (including massed or distributed practice and task scheduling). The article is useful for individuals concerned with the applicability findings of experimental data and theoretical concepts to practical learning situations.

Galanter, E. H. "Chairman's Summary of Conference." Abstracts presented at Air Force Office of Scientific Research and the University of Pennsylvania Conference on the Automatic Teaching of Verbal and Symbolic Skills. Philadelphia, December 1958. (Mimeo.) U

In this unpublished summary, not included in the book by Galanter, he reports conference agreement on these points: (a) the student learns something using automatic devices, (b) this is done in a reasonable time and perhaps with a real saving of time, (c) sequential variables in programming are critical, (d) steps in programming should be small. The areas which lacked agreement were: (a) the form of steps to be used, that is, the multiple-choice versus free response, and (b) error-rate, that is, should the ideal program contain no errors?

In summary, Galanter saw the teaching machine program and its structure to be the critical features of the program, even though some programs use no machine at all (such as book devices). He also noted that all of the discussion was on visual input materials. The use of auditory and tactual inputs was hardly mentioned probably because visual input materials are easier to construct; nevertheless, Galanter cautions that the importance of other types of input should not be overlooked.—*E. B. Fry et al.*

Galanter, E. H. (editor). *Automatic Teaching: The State of the Art.* New York: John Wiley & Sons, 1959. (a)

This is the first published book in the teaching-machine field, and is based on a symposium held in Philadelphia in December 1958, under joint sponsorship of the Air Force Office of Scientific Research and the University of Pennsylvania. Most of the articles included in the book are based on papers presented at the symposium, either comprising versions of the papers subsequently prepared by the authors or, in several instances, abstracts which they submitted in conjunction with the symposium. In addition, three papers not presented at the symposium are included: Galanter (1959), Gagné and Bolles (1959), and Pressey (1959). Abstracts of the articles in the book are included in the present Appendix. These include papers by Galanter (1959b), Gagné and Bolles, Beck, Skinner (1959b), Holland (1959), Meyer (1959), Porter (1959), D. E. P. Smith (1959), Homme and Glaser (1959b), Crowder (1959), Rath, Anderson and Brainerd, Briggs (1959d), Lumsdaine (1959a), Zeaman, Kendler, and Pressey (1959). Several reviews of this book are also noted in the Appendix: Barlow (1960b), Reid (1960), and Rothkopf (1960).

Galanter, E. H. "The Ideal Teacher." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 1, p. 1-12. (b)

This is an introduction to Galanter's book, which he wrote after the conclusion of the symposium on which the book is primarily based. Galanter takes the point of view that the design of a teaching machine is tantamount to the development of a theory of teaching. He surveys the problems represented by other papers in the book, and identifies the following as major issues emerging from the symposium. (a) The programming problem—what is the correct order of presentation of material? (b) The error rate problem—is there an optimum number of errors that should be made? (c) The step size problem—how far apart (in some sense) should adjacent items of the program be spaced? (d) The prompting problem—should prompts be used, and if so, how and what kind? (e) The constructive vs. multiple-choice response problems; (f) The pacing problem—experimenter-controlled rate vs. self-pacing; (g) The multiple-track problem—is one program satisfactory for all students? (h) The economic problem—how much is additional education worth? and (i) The generalization or testing problem—how do you know that the student has learned?"

He discusses the differences in rationale between Skinner's approach and Crowder's for taking account of the problems faced by the individual learner. He refers to backtracking and other forms of branching in a program, and also discusses the factors of pacing form of response, and ordering of items or frames in the construction of the program.

Gates, A. I. "Recitation as a Factor in Memorizing." *Archives of Psychology* 7: 40; 1917.

This well-known early study is of considerable interest because it is one of the pioneering investigations of active student response as a factor influencing the efficiency of learning. Gates compared the effectiveness of various ways of allocating the total available learning time between (a) merely reading over the material, and (b) actively attempting to recall or recite it. His results are summarized by Hovland (1951) as follows: "The results he [Gates] obtained with eighth-grade school children on the learning of nonsense syllables and short biographical prose selections show a pronounced advantage for recitation. . . . That the learning of nonsense syllables is affected more than the learning of prose has been confirmed by later investigators. Woodworth (1938) attributed the small advantage of recitation in prose to the fact that in rereading such material time after time one inevitably anticipates what is coming and consequently recites while reading."

Gilbert, T. F. "An Early Approximation of Principles of Analyzing and Revising Self-Instructional Programs." Paper read at the Air Force Office of Scientific Research and University of Pennsylvania Conference on the Automated Teaching of Verbal and Symbolic Skills, Philadelphia, December 1958. U (a)

This unpublished paper consisted primarily of an abbreviated exposition of the following report.

Gilbert, T. F. "An Early Approximation to Principles of Programming Continuous Discourse, Self-Instructional Materials." A Report to the Bell Telephone Laboratories. September 1958. (Unpublished paper, mimeo.) U (b)

This paper, though unpublished, is frequently quoted. Although Gilbert feels that much of its content is too provisional to warrant full publication, the Editors consider that it is worth abstracting in some detail.

Part I, the introduction to this report, includes two sections. Section A, *Behavior Principles and the Classroom*, presents principles concerning the reinforcement and shaping of operant responses that underlie classroom programmed learning. Section B, *Recent Attempts To Apply Behaviorial Principles to Teaching*, discusses very briefly the devices of Pressey, and somewhat more fully the Skinner devices and those of Porter and Ferster.

In introducing Part II, entitled *The Construction of Teaching Programs*, Gilbert cautions that his rules for writing teaching materials are only a crude approximation of optimal principles of programming. "These principles arise partly from the framework of behavior analysis discussed earlier, partly from the writer's experience in writing teaching programs, partly from observing neophytes learning to write programs, and partly from the experiences of others who have written programs. About all that can be said for their validity is that programs that were written according to them have been demonstrated to teach with considerable success and . . . [it] must be cautioned that, at this stage, the explication of these principles represents a highly personal document. It is an attempt by the writer to describe his behavior while writing programs. The behavior, even if accurately described, is not likely to be the last word in wisdom. It must be remembered that it is the essential problem of a research program in teaching to make a more exacting approximation to principles of programming."

Section A of Part II, *Designing Items To Evoke Relevant Behavior*, stresses the principle that a "subject matter should be analyzed into its stimulus and response components . . . . (a) The problem (or stimulus) that appears in the window of the automatic teaching device should represent the kind of situation to which you want to teach students to make responses. (b) The responses required of the student should represent the kind of response you want him to make."

Concerning principles of relevance, Gilbert points out that in his experience the most common tendency is to create irrelevant items in which the students are required "to make verbal descriptions of acts they are to learn rather than to write items that require the students to make the acts themselves." He continues: "The program writer has the further problem of arranging for contextual relevance. When teaching a student to respond to such problems as a + b = b + a as the commutative law for addition, the problem should also be stated in other symbols, such as k + m = m + k. In teaching trigonometry, a student is required to make certain descriptions of the characteristics of triangles. These triangles should be presented in various positions lest generalization fail to occur. In teaching the color code for resistors, supposing the resistors to be mounted on a breadboard next to the teaching device, the student should have to respond to resistors which are arranged in various positions, some projecting out from the board, some having to be read left to right, some right to left, and some up to down. Part of the analysis of subject matter will require the writer to state the class of relevant problems, rather than single instances, and provide a sufficient number of single instances so that responses to the class of stimuli can be made. In repeating items it is therefore well to repeat the items in different forms."

In Section B, *Problem Repetition*, Gilbert notes that: "It is desirable that any particular problems, or kinds of problems, should be repeated until the student's correct responses to them have achieved a high degree of probability of occurrence, and a minimum fading of this probability with time. Five conditions are desirable: (a) a minimum number of problems of a class; (b) a minimum sample of kinds of problems in a class; (c) a minimum time invested per student; (d) a maximum probability of correctess of response; and (e) a minimum fading of probability during a stated time. Particulars of a given class of problems, as particulars of any other class, will vary along several stimulus dimensions. The minimum number of necessary, particular problems will be at least two and probably a great number more. For example, the following factoring problems differ in several respects:

1. $A^2 + AB^2 = A(A + B^2)$
2. $7XY + 14X^3Y - 28XY^4 = 7XY\,(1 + 2X^2 - 4Y^3)$

These problems differ in the size of exponents, the number of terms, characteristics, notations, use of numbers as terms, and probably other characteristics. [If it] is desired that the student generalize the operation of finding prime factors for any set of notations, then at least two problems differing in notation should be presented. Repetition of problems should be different in as many particulars as possible. It seems plausible that repetition of problems should form a kind of density gradient, such that in the introduction of a new concept a number of problems are given close together and that later, repeated practices of these problems be spread more thinly throughout the program."

Gilbert points out the interesting notion that: "Certainly no exacting statements can be made about the numbers of repetitions required for certain kinds of problems. However, certain general guides can be suggested. For a particular type of material, some items will serve as building blocks for the accomplishment of more complicated material. For successful responses to the more complicated problems, less repetition of the more simple problems is necessary, since the simple problems will have to be solved over and over again in the solution of the more complicated ones." The amount of repetition needed in teaching paired-associate material can be reduced by introducing mediating associations, that can make the connection seem less arbitrary; for example, in teaching a color code for resistors, Gilbert provided a mediating association for each color-number pair, somewhat as follows:

ITEM	ANSWER
1. A black band on a resistor stands for 0: 0, black, nothingness. The black band stands for	0
2. A brown band on a resistor stands for the number 1: 1, brown, penny. A *black* band stands for	0
3. A number 1 is represented by a (a) .... band, a number 0 is represented by a (b) .... band.	(a) brown (b) black

Another principle of repetition is "dovetailing." This is appropriate where the student must learn to make a set of different responses to different, but related, stimuli. "For example in teaching algebra, we first introduce the commutative law for addition, following which the associative law for addition is introduced. After some repetition with the commutative law, the associative law is introduced with some repetition, and then the two laws are presented together, so that further repetition is done in such a way that the appropriate responses of the two sets of stimuli will not be confused." "In summary," Gilbert says, "we have found that the aversive characteristics of repetition can be reduced considerably by varying the context of each additional problem, by introducing new information related to a particular problem, and by requiring discriminations to be made between two or more problems separately accomplished in the past. . . . It may be necessary in the future to experiment with programs having varying amounts of repetition in order to discover some more exacting rules concerning this problem. It has also been suggested many times that it may be wise to write different programs for students at different levels of capacity. Thus bright students would not be burdened with excessive amounts of repetition."

Section C, *Augmenting and Fading*, discusses two general ways in which guidance through a program can be effected. One, *augmenting* involves introducing bits of information that can gradually lead a student to a new concept. The other method involves *fading* or "vanishing." [See Skinner's discussion on pages 146-47 of this volume.] In connection with the ordering of program sequences, Gilbert notes that one of the problems which comes up is the question of where to place statements that summarize the characteristics of a set of problems: "If a rule or generalization is given before the examples, it may in some way aid the student in working the exemplary problems. On the other hand, a series of examples may themselves be sufficient to lead the student to induce the general principle."

Section D, *Sequence of Machine Instruction, Expository Teaching, and Simulated Field Experience*, is concerned with the question of where self-instructional materials should be introduced within a total teaching situation—in relation to expository media (such as lectures or films), field experiences, etc.

Part III, *Suggested Procedures for Analyzing Continuous Discourse Teaching Programs*, includes three sections: Section A, *The Selection of a Continuous Discourse Program;* Section B, *The Characteristics of a Teaching Laboratory;* and Section C, *Suggested Schemes for the Analysis of Behavior*. In the last of these it is noted that three general kinds of information can be obtained about a teaching program: "First, the accuracy of the student in responding to items of the program, which can be measured by the count of item failure. Second, a measure of the rate at which the student goes through a program, which can be measured by the latency of response to an item or by the rate of cumulative responding to a series of items. Third, the ability of the student to generalize what he has learned in the teaching program to other kinds of situations."

Finally, Gilbert suggests a classification scheme in types of program "frames" or items according to their functions, using the following categories:

1. *Lead-In Items* neither give new information nor require the rehearsal of old skills. Their function is to orient the student to a problem and prepare him for new information.
2. *Augmenting Items* supply new information to a student but do not require him to make any relevant response. Responses required in such items may usually be of a kind that will help insure the students' reading the item.
3. *Interlocking Items* require a student to review established skills while new information is being presented.
4. *Rote Review Items* present a problem identical to one earlier presented. (They probably should be avoided except in the case where memorization of a short verbal statement is desired.)
5. *Restated-Review Items* require the rehearsal of a skill where the problem is restated in a different syntactical arrangement.
6. *Delayed Review Items* allow for further practice distributed in time. They will be different from other items only in time of presentation.
7. *Fading Items* require that the student review what has been presented to him; also, information is withdrawn from them from item to item.
8. *Generalizing Item* presents a verbal statement pointing to a common characteristic of several specific problems previously presented to the student.
9. *Specifying Items* exemplify a general rule in terms of specific cases. For example, after a student has been told that

multiplication should precede addition, he is then given problems such as: 2 + 3 × 9 = ________.

10. *Dovetailing Items* require the student to make separate responses to separable stimuli that otherwise become confused. For example, after teaching the appropriate number response to a red resistor and to a green resistor, a dovetailing item would require the student to make both responses to both colors.

Gilbert suggests that this scheme of classification or a similar one can be helpful in the analysis of a total program, and suggests that a program be analyzed in terms of these characteristics as it is being constructed. [Portions of this report are included in a later, published paper (Gilbert, 1959a).]

Gilbert, T. F. "Some Issues Involved in Selecting an Automatic Teaching Device." A Report to the Bell Telephone Laboratories, September 1958. (Unpublished paper, mimeo.) U (c)

This is a second widely cited article, which discusses the following topics:

A. *Presentation of Materials:*
   1. Variation in Frame Size;
   2. Modality of Presentation;
   3. The Mechanics of Presentation;
   4. The Type of Material Required; and
   5. Accessibility of Materials to Students.

B. *Reception of the Student's Response:*
   1. Reception of the Responses by Manual Operation;
   2. Handwritten Reception;
   3. The Reception of Vocal Responses; and
   4. Allowing the Student to Change an Answer.

C. *Confirmation:*
   1. Ways To Confirm the Correctness of the Response; and
   2. Giving Additional Information with Confirmation.

D. *Selection of Program Materials:*
   1. Discrimination Between Right and Wrong Responses;
   2. Retracing (program repetitions after a given number of failures);
   3. Sidetracking (shifting to additional and more complicated material after a string of successes);
   4. Skipping (termination of practice and skipping to the next topic); and
   5. Program Selection by the Student (the possibility of arranging an automatic teaching device so that the student can make selections that allow him to repeat certain materials, to get further information about the subject matter, or to skip on to later materials).

E. *Crowder's Scrambled Book Technique:* as compared to the Skinner approach [cf. Lumsdaine, 1960b].

Gilbert, T. F. "Some Recent Attempts at the Partial Automation of Teaching. A Report to the University of Georgia." 1958. (Unpublished paper, mimeo.) U (d)

This paper, like the two preceding reports by Gilbert, has received fairly wide circulation in mimeographed form even though it has not been formally published, and has been cited by several other writers. Gilbert gives an introductory discussion of behavior principles as applied to classroom teaching, employing primarily Skinnerian notions and discussing briefly the general role of reinforcement, the need for immediacy of reinforcement, setting up of differential responses by differential reinforcement and extinction, the relationship of reinforcement to motivation to work, and shaping of complex responses through successive approximations. He then discusses teaching machines more specifically and describes the characteristics of what he regards (or regarded at the time, at least) as an ideal machine. One of its features is that objective scoring of constructed responses is made possible through a coding scheme. Gilbert's description is as follows:

"The 'ideal' machine, which exists in only crude design, is a box about two feet square and nine inches high. A disk with a number of problems is placed in the machine. One problem is exposed under a glass window. In an open window there is exposed an inch or so of adding machine tape. The student writes on this tape what he believes to be the solution of the problem. He then pushes a button that advances his answer to a position underneath a glass window. He cannot retrieve his answer, nor can he erase it. The button that performs this operation also releases a lock on a device resembling an Address-O-Graph. This device is in the upper left-hand corner of the machine. A second button is pushed that opens the checking device to the appropriate place, exposing a set of 12 answers, one of which is correct. Each answer has a number. The student selects the number of the solution that matches his (if any do), and sets a selector dial to that number. At this time the checking device closes, and if the student's solution is correct a buzzer sounds and a new problem is presented in the problem window. Also, if the answer is correct, a hole is punched in the adding machine tape, recording the fact of a correct solution. If the student's solution is wrong, the tape remains in position, the problem remains in its window, and the checking device remains locked. He cannot proceed to a new problem until he has provided the right solution, so that he begins again and writes in a new answer.

"This machine allows the student to work at his own rate, and prevents him from randomly setting the selector dial to various positions until he hits upon the correct combination. These two features are provided for in the following way. Suppose that the machine has in it a number of problems in factoring algebraic expressions. A switching circuit in the machine is so designed that factoring problems keep appearing in the machine until the student has given five consecutive solutions on the first attempt. The chance of obtaining five consecutive correct selector switch settings on the first trial is, in this case, about one in 250,000. A student who tried to beat the machine by making random selector switch settings would have factor problems repeated to him during the allotted time. When the student actually comes to understand the principles of factoring well enough to give five consecutive first trial answers, the machine automatically drops out the factoring problems and begins to present the next higher level material. The machine literally 'senses' the fact that the student has induced the principle of factoring."

Gilbert also gives a few very brief remarks on the problems of writing programs; his notions on programming are expanded in much greater detal in his reports to the Bell Telephone Laboratories.

Gilbert, T. F. "An Early Approximation to Principles of Programming Continuous-Discourse, Self-Instructional Materials." *Psychological Problems and Research Methods in Mathematics Training.* (Edited by R. L. Feierabend and P. H. DuBois.) Washington University, USOE Cooperative Research Project No. 642, June 1959. (Mimeo.) L (a)

This is a published version incorporating materials from the author's earlier reports (1958 a, b, and c).

**Gilbert, T. F. "On the Relevance of Laboratory Investigation of Learning to Self-Instructional Programming," Symposium paper presented at the annual convention of the American Psychological Association, Cincinnati, September 1959. [p. 475-85]. (b)

**Glaser, R. "Christmas Past, Present, and Future." *Contemporary Psychology* 5: 24-28; 1960. [p. 23-31] (a)

Glaser, R. "Comments on the Use of Teaching Machine Techniques in the Language Laboratory." *Language Teaching Today* (Edited by F. J. Oinas.) A report of the Language Laboratory Conference held at Indiana University, January 22-23, 1960. Publication No. 14 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics. (In press 1960). (b)

This paper emphasizes the importance of programming and program construction, as contrasted with development of equipment for use in language laboratories. The basic importance of well-programmed subject material is reiterated as the keystone of effective language-laboratory instruction.

Glaser, R. "Principles and Problems in the Preparation of Programmed Learning Sequences." A paper read at the University of Texas Symposium on the Automation of Instruction, University of Texas, May 1960. [Also to be published as a report of a Cooperative Research Program grant to the University of Pittsburgh under sponsorship of the U. S. Office of Education. In press, 1960.] (c)

This paper is a review and summary of key points in the preparation of programmed learning sequences and their use in teaching machines as described by the major papers in existence as of 1960. The topics commented on are the following: defining the subject-matter field; the central role of reinforcement; gradual progression to establish complex repertoires; step size and difficulty; prompting and increasing the probability of the desired behavior; the principles of fading and vanishing; confirming mechanisms and scoring; the teaching of observing or attending behavior; practice and review; teaching understanding and discrimination; inductive and deductive programs; editing and revising a program; weaning from the machine; adaptive programming; response modes; individual differences; subject-matter characteristics; machine modes; and control of the learning process and experimental investigation.

**Glaser, R.; Damrin, D. E.; and Gardner, F. M. "The Tab Item: A Technique for the Measurement of Proficiency in Diagnostic Problem Solving Tasks." *Educational and Psychological Measurement* 14:283-93; Summer 1954. [p. 275-85]

Glaser, R., and Glanzer, M. *Training and Training Research.* Pittsburgh, Pa.: American Institute for Research, August 1958. △ L

This report considers the following aspects of the training process; (a) The specification of training objectives; (b) Input control, the selection of personnel for entrance into training; (c) Training methodology; and (d) Output control, the measurement of proficiency attained as a result of training. For each of these aspects, contributions from psychological research are considered and research problems are defined. Section Two, on training methodology, examines the principles of learning which are relevant to the practical trainer and educator and are also relevant for the setting up of training research problems.

**Glaser, R.; Homme, L. E.; and Evans, J. *An Evaluation of Textbooks in Terms of Learning Principles.* Paper presented at the meeting of the American Educational Research Association Atlantic City, N. J., February 1959 [p. 437-45]

Glaser, R., and Phillips, J. C. *Guided Missile Personnel Research: Report No. 5, An Analysis of Tests of Proficiency for Guided Missile Personnel: II. The Troubleshooting Board.* Pittsburgh, Pa.: American Institute for Research, 1954.

This Technical Report to the Office of Naval Research describes an analysis of a tryout administration of the Troubleshooting Board which is a formboard version of the Tab Item. [See Glaser *et al.*, p. 275-85.] Test items are set into a frame, and answer feedback is obtained when the subject lifts up a block to designate his response.

Goldstein, M., and Chalmers, E. L., Jr. "Photo-Identification Concepts via Teaching Machine." Unpublished paper. (1960) U

This as yet unpublished paper, prepared for Thompson Ramo-Wooldridge, represents an interesting application of the teaching-machine approach to a highly specialized problem encountered in military intelligence work. (Goldstein is now at Princeton University and Chalmers is at Florida State University.) The authors derive a classification, and describe a number of classes based on the relation between stimulus properties and the kinds of generalization which the learner is required to make. In the light of this analysis, a paradigm is outlined that identifies requirements for a teaching machine which would seemingly be adapted to cope with teaching the kinds of concept formation involved in this task. The authors then examine briefly the differing approaches of Skinner and Crowder with respect to the control of errors in practice, and propose a program-sequencing arrangement which they describe as a compromise. They also discuss the design of teaching-testing cycles in relation to setting up of discriminations to develop necessary concepts.

Grings, W. W.; Rigney, J. W.; Bond, N. A.; and Summers, S. A. *A Methodological Study of Electronics Troubleshooting Skill: I. Rationale for and Description of the Multiple-Alternative Symbolic Troubleshooting Test.* Technical report prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, August 1953. (Mimeo.) △ L

This report concerns a unique type of test format which was devised in studying electronic troubleshooting. The report gives a detailed description of the test format, which is the so-called "MASTS" test. This test, like the "tab test" described by Glaser *et al.* (1954), permits contingent sequences of responses to be selected and scored. This is accomplished by a relatively simple apparatus which does not involve complex electrical or electronic circuitry.

Hatch, R. S. "An Evaluation of the Effectiveness of a Self-Tutoring Approach Applied to Pilot Training." Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, 1959. WADC Technical Report 59-320. △

A Navy Automatic Rater multiple-choice card teaching machine programmed with items on instrument flying was made available to one squadron in its lounge for two months. Pre- and post-testing compared with a control squadron showed a significant gain in job-information for a squadron having the incidental use of the automatic rater.—*E. B. Fry, et al.*

Hively, W. "Implications for the Classroom of B. F. Skinner's 'Analysis of Behavior.'" *Harvard Educational Review* 29: 37-42; Winter 1959.

Hively gives a theoretical discussion of Skinner's principles of learning. Those cited are: (a) there is need for the rewarding of successive approximations, (b) immediacy of reward is very important, (c) high frequency of reinforcement is important, (d) scheduling of reinforcers can control frequency of response, (e) negative reinforcers are bad because of elicitation of side effects, and avoidance which may be learned rather than the task.

Other principles mentioned are the need to isolate the subject and the use of a machine in the classroom for adequate reinforcement. With respect to teaching machines he says: "They differ mainly in form of response required—multiple-choice, the construction of an answer from a limited number of letters or numerals, or a written answer which is confirmed by comparison with the correct answer. Each appears to be particularly adaptable to certain subject matter and age levels, but all have in common the presentation of a sequence of problems and provisions for immediate confirmation of response to each problem." (p. 40)—*E. B. Fry, et al.*

**Hively, W. "An Exploratory Investigation of an Apparatus for Studying and Teaching Visual Discrimination, Using Pre-school Children." (1960) [p. 247-56]

Hoehn, A. J., and Lumsdaine, A. A. *Design and Use of Job Aids for Communicating Technical Information.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, January 1958. Technical Report AFPTRC-TR-58-7. (ASTIA Document No. 152109.) △

In addition to reporting studies on the use of job aids for more effective performance in troubleshooting, this report briefly describes equipment for automated procedural training (see Lumsdaine, p. 17-19). It also describes a series of studies that investigated the comparative effectiveness of different ways of using filmed demonstrations alternated with active practice (see also Maccoby and Sheffield, 1958).

Holland, J. G. "A Teaching Machine Program in Psychology." *Auto-*

matic Teaching: The State of the Art (Edited by E. H. Galanter). New York: John Wiley & Sons, 1959. Chap. 5, p. 69-82.

In this study, 187 sophomores used Skinner's disk-type teaching machines for 10 weeks. The subject matter was related to Skinner's book, *Science and Behavior*. The students worked for a median time of 14 hours to cover some 48 disks of 29 frames each; this meant a total of 1400 different frames. The time to work through the easiest disk ranged from three to 14 minutes, the average being 10 minutes. On disks of medium difficulty the median time was 14 minutes; on the most difficult, 28 minutes. Although students were permitted to distribute the work themselves, they never took more than six disks per session. Seventy-six percent of the students said that the machine helped them in studying. An analysis of performance on the final examination showed that items on which the machine could have been expected to help were much more frequently answered correctly than were items for which the machine work was irrelevant. These are some of the programming techniques used by Holland:

1. Panels consisting of short passages of printed material or graphs to which the student could refer at any time while working the item in the machine.
2. The use of new words in a series of frames before requiring the student to use them.
3. Some items moving from definition to examples within a single frame.
4. Some items using "leading-in" from an assumed knowledge.
5. The indication of relevant categories for response terms.
6. Some items using high association words or common phrases.
7. Avoidance of undesired alternative responses by careful phrasing of the frames.

The strengths and weaknesses of the program were revealed by an item analysis of student responses.—*E. B. Fry, et al.*

**Holland, J. G. "Teaching Machines: An Application of Principles from the Laboratory." (1960) [p. 215-28]. [Also published in *Proceedings of the Educational Testing Service Invitational Conference*, October 1959, on "The Impact of Testing on the Educational Process." Princeton, N. J.: Educational Testing Service, 1960. (In press)]

Holland, J. G., and Skinner, B. F. "A Self-Tutoring Introduction to a Science of Behavior." Unpublished manuscript. (Cambridge, Mass.: Psychological Laboratories, Harvard University, Circa 1958-59.) U

This is a teaching-machine program used by Holland and Skinner in teaching a behavioral science course at Harvard.

**Homme, L. E. "The Rationale of Teaching by Skinner's Machines." Paper read at the annual convention of the American Psychological Association, September 1957.

**Homme, L. E., and Glaser, R. "Problems in Programming Verbal Learning Sequences." Paper read at Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Convention, Cincinnati, September 1959. [p. 486-96] (a)

Homme, L. E., and Glaser, R. "Relationships Between the Programmed Textbook and Teaching Machines." *Automatic Teaching: The State of the Art* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 9, p. 103-108. (b)

The topic of this paper is elaborated upon in more detail in the papers by Glaser, Homme, and Evans [p. 437-45] and by Evans, Glaser, and Homme [p. 446-51]

Hovland, C. I.; Lumsdaine, A. A.; and Sheffield, F. D. *Experiments on Mass Communication*. Princeton, N. J.: Princeton University Press, 1949.

This book reports on experimental studies of the effectiveness of training in indoctrination films conducted during World War II. In addition to the specific studies reported, it contains several technical appendices dealing with problems in experimentation on instructional media which have applicability to experiments on the effects of self-instructional materials. Of the experimental studies reported, the training experiment described in Chapter 9 is the one which deals most directly with procedures closely similar to those employed in teaching-machine instruction. Graphs presenting some of the major findings are also included in the paper by Lumsdaine (1959a) in the book edited by Galanter (1959a). In this study a film was used to teach the military phonetic alphabet, A for able, B for baker, etc. Two forms of the film were compared in order to study the effect of using active student response procedures (so-called "audience participation") during review sequences. The difference between the two experimentally-compared films was solely in the review sequences. The control film used a standard or "passive" form of review in which letters were presented along with their phonetic equivalents. In the active review version, the letters only were presented, and the students were instructed to try to call out the correct equivalents for each letter in turn. The "prompts" thus provided for the students' responses lay in the fact that correct equivalents had been recently presented, providing a basis for their being emitted during the review. Feedback-correction and/or further "prompts" were provided by the fact that the group had to call out each response aloud. Since someone almost always gave a recognizably correct answer, correction-prompts were thereby provided. The superiority of the active-response procedure over the control procedure was highly significant. The results also showed that the advantage of the active-response and feedback procedure was least when it might be considered least needed and most when it would appear most needed—that is, greatest for less motivated, slower students learning more difficult material, and least for brighter, highly motivated students in learning the easier material.

Irion, A. L. "An Investigation of Four Modes of Operation and Three Types of Learning Tasks on the Improved Subject Matter Trainer." Unpublished paper, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., 1957.

This is a preliminary version of the following paper by Irion and Briggs.

Irion, A. L., and Briggs, L. J. *Learning Task and Mode of Operation Variables in Use of the Subject Matter Trainer.* Lackland Air Force Base, Texas: Air Force Personnel and Training Research Center, October 1957. Technical Report AFPTRC-TR-57-8. (ASTIA Document No. AD 134252) △

This experiment uses the Subject Matter Trainer described by Briggs, 1958, together with its modes of operation. In this experiment the modes varied in effectiveness in descending order, as follows: (a) quiz mode, (b) modified quiz mode, (c) practice mode, and (d) single-try mode. In the quiz mode the student quizzes the machine. In other words, he merely presses a button which always shows the right answer and is instructed to attempt to remember the right answer; he then presses another button which causes the machine to proceed to the next item. The same mode is referred to by Briggs, (1958), as the "coaching mode." In the modified quiz mode, second in effectiveness, the student is told the correct answer after the first try. Briggs referred to this as the "single-error permitted mode."

The authors account for the superiority of the quiz mode as a result of the following four factors:

1. No time is lost searching for the correct response, so more trials per time unit are accomplished.
2. Overt wrong responses are avoided, although covert wrong responses are possible if the student fails to follow directions precisely.
3. No signal, buzzer, or red light for a wrong response is ever received.
4. There is a close temporal continuity between presentation of the stimulus and the appearance of the green light.

This experiment did not compare learning rates on this and on more conventional devices. However, the authors did state that the modes of operation of the Subject Matter Trainer were "surprisingly effective," and that the results of the study appeared to justify further use of the Subject Matter Trainer. However, they suggest that a simplified model could be used for operational use; this modified model would handle only the quiz and modified quiz modes of operation.—*E. B. Fry et al.*

Israel, M. L. "Educational Redesign: A Proposal." *School Review* 67:292-304; 1959.

"Dissatisfaction with the accomplishments of our schools is widespread. One difficulty seems to be that the expanding aims of the modern educational system have far outdistanced its present capabilities for shaping behavior." Education is viewed as an engineering branch of psychology with the educator molding behavior in the class room in accordance with the principles developed for the psychologist in the laboratory. Based on the notions expressed in Skinner's *Science and Human Behavior,* suggestions are made for increasing the effectiveness of available sources of the educational control of behavior. The question is raised as to whether the educator can be sure that appropriate transfer will take place once the individual is outside the educational environment. It is suggested that it might be wise to allow the educator to supervise the transfer from educational control to the maintenance of similar behavior in society.

Israel, M. L. "Variably Blurred Prompting: I. Methodology and Application to the Analysis of Paired-Associate Learning." *Journal of Psychology* 50:43-52; 1960.

This is an experimental study using a technique of "partial prompting" through the use of an indistinct image. This technique of using a blurred cue to achieve partial and variable prompting is similar to the method described by Lumsdaine [p. 537] in a theoretical discussion of optimal cue-strength gradients. The learning materials used by Israel consisted of English words for the response terms and, in the experimental trials, of German words for the stimulus terms. The purpose of the study was to obtain evidence for the existence of partial associations by comparing each subject's performance on experimental trials, in which the blurred cue (English word) was used as a prompt in paired-associate learning versus control trials in which the blurred cue was also used but with no German word present as a stimulus term.

Whenever the subject could not recall the English word corresponding to a German word (experimental trials), he operated a control which caused a blurred image of the English response term to be exposed on a movable screen, which he then moved to bring the image gradually into clearer focus. In the control trials the subject's task was simply to bring the blurred image of the English word into sufficiently clear focus that he could barely recognize it. (Israel does not report any direct comparisons between learning with the variably indistinct cue and learning under direct prompting without the variable clarity cue.) The amount of clarification required was less in the experimental condition than in the control condition. The author concludes that "results demonstrate the existence of partial associations, but do not show clear evidence for a growth in their strength prior to the point at which unaided recall is possible."

Jensen, B. T. "An Independent-Study Laboratory Using Self-Scoring Tests." *Journal of Educational Research* 43:134-37; October 1949.

Two groups totaling 24 mature and superior students completed a five-hour course in psychology in five weeks in an independent-study laboratory, 54 percent obtaining "A" grades on the basis of two objective midterms and a final examination, as compared to 10 percent so graded in 27 sections of the course taught in the usual manner. A distinctive feature of the laboratory was an extensive set of tests covering the entire course, used with the "punchboard" so as to make them self-instructional. These tests were also used as the point of departure for selective review and discussion with other students or the assistant in charge, when needed. There was no teaching in the usual sense. The students not only made good grades, but reported gains in capacity for independent work, and saved time so as to take an extra course or engage in some other undertaking not otherwise feasible.—*S. L. Pressey.*

Jones, H. L., and Sawyer, M. O. "A New Evaluation Instrument." *Journal of Educational Research* 42:381-85; 1949.

These authors report an experiment to assess the use of a punchboard patterned after the one described by Angell and Troyer for providing self-scoring and immediate knowledge of results in weekly quizzes given to ten sections of college freshmen in a course in citizenship. Half of the classes (one per instructor) used the punchboard for each of eight weekly quizzes. The other group of classes used conventional answer sheets that did not provide knowledge of results of self-scoring. Eighty-three percent of the experimental group said they preferred the punchboard method, though a considerable number expressed various reservations such as that there was no way to erase an answer on second thought. Mean scores on a final examination were 23.7 for the experimental group and 21.7 for the control group, with a *t* ratio of 1.7 based on N's of 116 and 112.

Jones, R. S. "Integration of Instructional with Self-Scoring Measuring Procedures." *Abstracts of Doctoral Dissertations* 65: 157-65; 1954. (Columbus: The Ohio State University Press.)

This study basically confirms Pressey's 1950 study which emphasizes the desirability of using the punchboard. Jones thinks that the most valuable instruction techniques were "individualizing instruction, the provision of special opportunities for the extremes of the groups, the use of self-appraisal, and diagnostic teaching." He believes that punchboards reach this goal better than do "conventional teaching" methods.

Jones commented on the programming of multiple-choice materials, declaring that items which were especially selected for learning value were better than the typical test-type items. Good learning items are "applicational, pose a definite problem, call more for thinking than for just memorizing, and are between .40 and .60 in difficulty."—*E. B. Fry et al.*

**Keislar, E. "The Development of Understanding in Arithmetic by a Teaching Machine." *Journal of Education Psychology* 50:247-53; 1959. [p. 425-36] (a)

Keislar, E. "Theoretical Aspects of Automated Teaching." Paper read at the annual convention of the Western Psychological Association, San Diego, Calif., April 1959. U (b)

Keislar states that we do not have a theoretical system for automated teaching. Rather, we have a collection of principles or working hypotheses drawn largely from reinforcement theory. He suggests two other approaches, both of which are being developed: (a) information theory, and (b) game theory.

He cites two types of machines which seem to be at theoretical extremes. The first is a "successive approximations" model developed by Skinner. In this type of machine, the stimulus has a high probability of eliciting the correct response. The second stimulus can very easily be generalized from the preceding item; hence, it also has a high probability of response. Each successive step leads gradually to a predetermined goal.

The second type of machine, which Keislar calls "Ideal," would require a computer with a large storage mechanism. Its chief characteristic is the modification of the program based upon the performance of the learner. In short, the program is adapted to individual differences. Such a device is geared to skipping steps for rapid learners, as well as to remedial and enrichment programs. In the ultimate machine, each step would be dependent upon everything that had gone before. This ultimate machine could have two types of goals: one an objective, instructor-determined, fixed goal; the other determined, at least in part, by the student's preceding performance. Hence we see that even the goal itself may be determined after the student has started performing on the machine.

Keislar is also concerned with some of the problems involved in programming. Stimulus generalizations and response generalizations are fairly obvious. But one thing that must be taken into account in human learning is a mediating generalization. Sometimes this mediating generalization may be planned within the program; usually, however, it would involve prior learning that had occurred elsewhere. The program constructor must, therefore, rely on his own information about the student's previous learning. Since this is little understood, programming is still largely an "art." Keislar mentions two criteria that may be applied: (a) Is the interval between steps small enough to hold the probability of a correct response high? (b) Is the interval large enough to facilitate rapid progress?

Keislar identifies two other problems confronting the program writer: (a) Where to begin? (b) What route or sequence to follow? Though there are no set answers to these problems, automated teaching provides at least an empirical test of them.

Kendler, H. H. "Teaching Machines and Psychological Theory." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 15, p. 177-86.

Skinner's advocacy of the teaching machine as an educational instrument seems eminently reasonable. The teaching machine holds promise of being the most efficient exploiter of a powerful learning mechanism, viz., immediate reinforcement (feedback). The teaching machine is also an efficient device to shape a student's behavior so that he may gradually proceed from simple concrete ideas to complex abstract ones without being retarded or pushed ahead too rapidly by his fellow students.

Although the recent development of the teaching machine is obviously an expression of Skinner's interest in controlling behavior, it would be both useful and advantageous to distinguish between the teaching machine as an educational machine and as an expression of Skinner's atheoretical position in the psychology of learning. The purpose of this paper is to raise problems associated with the teaching machine that seem to demand the collaboration of theoretically oriented psychologists.

According to the S-R mediational theory which has been developed in somewhat various forms by different investigators much of the behavior that occurs in school can best be accounted for by assuming certain events intervene between the stimulus and response. By understanding the principles that operate in the student's ability to generate implicit responses that function as cues for overt behavior the educator can best utilize the teaching machine. Recent work in problem solving and language behavior suggests that such factors as (a) language ability, (b) concept formation and concept utilization, (c) transfer, (d) individual differences, and (e) the criterion of learning be considered in the programming of teaching machines. —*H. H. Kendler.*

Kimble, G. A., and Wulff, J. J. *The Effect of "Response Guidance" on the Value of Audience Participation in Training Film Instruction.* Washington, D. C.: Air Force Human Factors Operations Research Laboratories, 1953. (HFORL Report No. 34.) △ L

Kimble and Wulff were interested in what is basically the "crutch" or "prompt" question—that is, the amount of guidance to be provided as students proceeded through a practice schedule that alternated short practice exercises with short segments of exposition or demonstration. Automated exposition, to teach the reading of logarithmic scales employed in the slide rule and similar analog computing devices, was presented to groups of subjects in a classroom situation by use of a filmstrip accompanied by a synchronized tape recording. Students did practice exercises in workbooks in which slide-rule scales or portions of them were reproduced. Responses consisted of giving the numerical value of a marked scale reading, or, conversely, finding a specified value on the scale. All groups were given the same exposition and the same amount of practice. The major experimental variable was the use of prompts or response guidance. The guided-practice group was provided with cues which limited the possible responses but without specifying exactly the correct response. The no-guidance group was allowed to make more errors, by withholding such prompts. (In all cases the students had already been told how to do a similar exercise and that further instruction was given as follow-up after each exercise.) The results showed a clear margin of superiority for the prompted group, which did better, on a later test, not only on the items used in practice but also in transferring to other similar items. The margin for superiority was greater on hard items than on easy ones, and was greatest where the prompts had been chosen to help the student avoid certain types of errors which preliminary work had shown were the ones most commonly made.

Klaus, D. J. (in collaboration with Holmok, E. S.; Lumsdaine, A. A.; Meyer, D. S.; and Quisenberry, K. S.) *High School Physics.* (A two-volume programmed text: Vol. 1, *Introduction and Static Electricity,* 545+ p.; Vol. 2, *Reflection and Refraction,* 720+ p.) Pittsburgh, Pa.: American Institute for Research, 1960. L

This program covers approximately six weeks of high-school physics on the topics of electricity and optics. Though published, it is of limited availability since most of the approximately 500 copies printed were utilized in the field study for which the materials were prepared (see Klaus and Lumsdaine, 1960). The program contains a total of 2928 frames, distributed as follows: Electricity I—Static Electricity and Direct Current—nine units comprising 1204 frames; Light I—Nature and Reflection of Light—six units comprising 826 frames; and Light II—Refraction of Light—six units comprising 809 frames. The first volume also includes an 89-frame refresher unit on mathematics. As prepared for field work accomplished up to the date of publication of this book, the self-tutoring frames are presented in "programmed text" format. (See similar format described by Glaser, Homme, and Evans, 1959, as illustrated on pages 437-45).

The programmed materials are also designed to permit later reproduction on microfilm with a suitable teaching machine. The programmed materials were evolved through a process of successive tryout and revision, in which the responses of trial students were used as a basis for frame revision. Numerous pictorial and diagrammatic illustrations are used in many of the frames; also, in other frames the response required of the student is to construct or draw or complete a diagram.

Klaus, D. J., and Glaser, R. *Studies of Navy Guided Missile Teams.* Pittsburgh: American Institute for Research, December 1958. △ L

This report consists of a brief summary of activities conducted in a program of training research for the Office of Naval Research. It includes a digest of project activities, a short administrative history of the project, and an annotated bibliography of the reports produced. The series of interrelated studies concern the specification of training objectives, the development and evaluation of the Troubleshooting Board, which is a variation of the tab item procedure (Glaser, Damrin, and Gardner, 1954), and the study of individual and team training procedures.

Klaus, D. J., and Lumsdaine, A. A. *An Experimental Field Test of the Value of Self-Tutoring Materials in High School Physics.* An Interim Report of Progress and Preliminary Findings. Pittsburgh, Pa.: American Institute for Research, April 1960. (Mimeo.) L

This report describes the progress in development of self-tutoring materials and preliminary findings from initial field tests, obtained in a study conducted by the American Institute for Research. This study is sponsored by the United States Office of Education under Title VII of the National Defense Education Act, to assess the effectiveness of applying recently developed self-instructional techniques in ongoing instruction. Its specific objective is to determine the contribution made to achievement in high-school physics by use of materials employing self-tutoring techniques. (See Klaus, 1960.)

Preliminary results reported are based on comprehensive examinations covering the first nine units of the self-instructional material. These results include test data from approximately 450 students in 15 schools in Western Pennsylvania which are regularly utilizing the Harvey White physics films, telecast daily over stations WQED and WQEX in Pittsburgh. Comparison of results for experimental classes to whom the self-instructional materials were made available to supplement the physics course (as compared with control classes who did *not* receive the self-tutoring materials) demonstrate that the self-tutoring materials have made a significant contribution to the level of achievement attained by participating students. The magnitude of the increase in achievement is great enough to conclude that, when properly written, self-instructional materials represent an encouraging development in educational procedure. Further plans under this project are also summarized in the report.

Komoski, K. P. "Origin of Collegiate School Automated Teaching Project." Unpublished report. New York: Collegiate School, 241 W. 77th St., February 1960. U

A two-year project was started in November 1959 with a grant from The Fund for the Advancement of Education. Rheem Califone supplied 15 write-in machines and IBM made a computer laboratory available for analysis of data. Programming is in progress in the following areas: (a) *French Spelling* for the sixth and seventh grades to supplement the direct (conversation) method. "Preliminary studies seem to indicate that this program is approximately twice as effective as other spelling methods." (b) *Pre-Verbal Number Concepts.* (c) *Elementary Arithmetic;* a revision of Susan Meyer's program. (d) *Sets, Functions, and Relations* for high school as recommended by the College Entrance Examination Board Commission on Mathematics—this program has been put on programmed texts available for 30 eighth-graders. (e) *Elementary Statistics* programmed text written by Evans, Homme, Glaser, and Stetler will be used on twelfth-graders. (f) *Remedial Spelling* for high school students based on 100 words commonly misspelled by college freshmen. Future plans of the project include further programming, experimentation, and teacher's institutes.—*E. B. Fry et al.*

Kopstein, F. F., and Roshal, S. M. "Method of Presenting Word Pairs as a Factor in Foreign Vocabulary Learning." Paper read at the annual convention of the American Psychological Association, San Francisco, Calif., September 1955. U

The influence of stimulus and response term synchronization on the learning of a list of eight foreign words, presented as paired associates with their English equivalents, was studied. Under condition "A" the English stimulus term and the foreign response term of each item were projected on a screen simultaneously for 3½ seconds; under condition "B" the stimulus term alone was displayed for the first 2 seconds and was then joined by the response term for the remaining 1½ seconds. Twelve repetitions of the eight-item pair list were presented. The fourth, eighth, and twelfth repetitions were used as test series. The material was presented to 48 groups of basic airmen under the first condition and to another 48 groups under the second condition.

Contiguity learning theory suggests advantages for the simultaneous presentation of both terms, and also suggests that the superiority of one method over the other might be a function of the point on the learning curve at which the measurement was made.

As predicted, the results indicate that the simultaneous presentation method is significantly superior in the early stages of training. This advantage for the simultaneous presentation method becomes distinctly less important as more substantial levels of learning are reached.

**Little, J. K. "Results of Use of Machines for Testing and for Drill upon Learning in Educational Psychology." *Journal of Experimental Education* 3:45-49; September 1934. [p. 59-65]

Lumsdaine, A. A. "Possibilities and Implications of Automated Teaching Methods." Paper read at the annual convention of the American Psychological Association, Washington, D. C., August 1958. U

This was an introductory paper read at a symposium on "The Automation of Teaching." Symposium participants were B. F. Skinner, E. Z. Rothkopf, N. A. Crowder, S. L. Pressey, and D. Zeaman. [See separate entries concerning these papers.] A parallel was drawn between the impact of automation on other fields and its potential application in education. In contrast to other forms of semi-automation represented by photographic or electronic dissemination of lectures and demonstrations through educational films and television, emphasis was placed on the teaching of individual students by teaching machines. Three devices were illustrated which had been developed under the AFPTRC program in Denver. These devices included the Subject Matter Trainer (see Briggs, 1958), a magazine-loading film projector for step-by-step demonstrations of procedures (see Lumsdaine, 1959d) and a variable-program microfilm device (see Crowder, 1960b, Lumsdaine, 1959d).

Lumsdaine, A. A. "Partial and More Complete Automation of Teaching in Group and Individual Learning Situations." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 13, p. 147-66. (a)

In this paper it is suggested that automated instructional methods comprise, generically, any means whereby teacher or tutor functions are provided by instructional sequences that are prepared in advance and are capable of instructing effectively without direct intervention or modification by a teacher. They may thus include not only individual teaching machines as such, but also certain forms of recorded instruction by films, tapes, etc. The features of active response and prompt knowledge of results are not found in the usual lecture or its automated surrogate represented by customary forms of films and instructional television presentation, but have been achieved with special kinds of group-presented films or recorded instruction that provides for frequent interspersing of questions between segments of presentation. Examples developed in the Air Force for technical training and language instruction are illustrated. Past research, on factors influencing the effectiveness of active-response procedures in conjunction with instructional film presentation, has applicability to the problems of teaching-machine programming. This was illustrated by several studies: see separate abstracts for Hovland, Lumsdaine, and Sheffield, 1949; Michael and Maccoby, 1953; Kimble and Wulff, 1953; and Maccoby and Sheffield, 1956.

Lumsdaine, A. A. "Response Cueing and Size-of-Step in Automated Learning Programs." Paper read at Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Convention, Cincinnati, Ohio, September 1959. U (b)

This was a preliminary version of the following paper.

**Lumsdaine, A. A. "Some Issues Concerning Devices and Programs for Automated Learning." (1959) [p. 517-39] (c)

**Lumsdaine, A. A. "Teaching Machines and Self-Instructional Materials." *Audio-Visual Communication Review* 7:163-81; Summer 1959. [Reprinted in this volume, with minor changes, under the title "Teaching Machines: An Introductory Overview" [p. 5-22]. A condensation of this paper also appeared in *Education Digest* 25: 1-4; December 1959. (d)

Lumsdaine, A. A. "Design of Training Aids and Devices." *Human Factors Methods for Systems Design.* (Edited by J. D. Folley, Jr.) Pittsburgh: American Institute for Research, 1960. L (a)

This paper discusses the requirements, functions, and characteristics of training devices, with particular but not exclusive reference to problems of military training, for achieving five classes of training objectives: identifying, knowing principles and relationships, following procedures, making decisions, and performing skilled perceptual-motor actions. Representative devices are illustrated and described in relation to characteristics that appear important for achieving these various objectives. A number of the illustrative devices described were developed at the former AFPTRC Maintenance Laboratory. The disadvantages of "mixed-objective" devices, particularly those used merely as visual aids for a classroom-lecture presentation, are compared with the advantages of devices orientated toward the attainment of specific objectives and designed to provide suitably guided individual student practice. A suggested outline of procedures for the design of training devices is presented.

Lumsdaine, A. A. "The Development of Teaching Machines and Programmed Self-Instruction." *New Teaching Aids for the Amer-*

ican Classroom. Stanford University, Calif.: The Institute for Communication Research, 1960. p. 136-73. (b)

This paper, together with others in the same volume, was originally presented in November 1959 to a symposium at the Center for Advanced Study in the Behavioral Sciences, under sponsorship of the United States Office of Education. It is a review and assessment of the current state of the art in research and development on teaching machines and individual programmed self-instruction. It was pointed out that "automation" of teaching processes, represented by these developments or other forms of automation, do not mean removing the process automated from intelligent human control—but, rather, merely that the control is transformed or encoded so that it can be displaced in time and space and indefinitely reduplicated in a standardized way. It was noted that, as in the case of the digital computer, the only intelligence which an automated teaching system can display is that which has been designed or programmed into it by man.

The essential difference between automation in teaching machines and that exemplified in television and instructional films lies not in the physical configuration of the devices but rather in the kind of teaching function which it is thought to automate or simulate: Teaching machines automate the tutorial process, whereas the other media tend to automate group instruction in the lecture-demonstration form. Teaching machines and programmed self-instructional materials seek to provide the advantages of the tutorial method in a way that is economically feasible. An extensive review is given of various kinds of teaching machines and the programming of verbal subject matter is discussed in some detail. Alternative approaches to programming and assumptions underlying these are pointed out. Current program construction, involving a great deal of ingenuity and skill, still remains more of an art than a science. Steps in programming include: (a) Identify as exactly as possible the behavioral repertoire to be achieved; (b) Analyze the desired behavioral outcomes into elementary components; (c) Identify logical and sequential interdependencies of elements as a basis for sequencing the instruction so that each step goes on with what has preceded it; (d) Devise frames for eliciting each elementary unit of heuristic or terminal behavior desired in proper context and strengthening it a unit at a time, using appropriate prompting techniques; (e) Compose a tentative program based on the foregoing; (f) Try out the trial program with a typical student, revising it until the student's responses show that he is really mastering every step; (g) Organize the program into units, using such optional or contingent sequences as are necessary to adapt to individual differences; (h) Recast the format of the program to fit the capabilities of machines which are available or can be designed to present it; (i) Utilize the resources provided by a teacher to augment the flexibility of the program in handling the contingencies, redirecting the student, assigning remedial programs or working individually on special aspects which it has not been economically feasible to program fully.

A review of device characteristics and of issues for research and development underlying the design of teaching machines and self-instructional programs is included; this parallels the discussion and includes the pictorial and tabular material contained in the earlier paper by Lumsdaine (1959c) [See pages 517-39.]. An analysis of the programming approaches used by Crowder, on the one hand, and by Skinner and his followers on the other, is given in the following terms:

"The two approaches have in common the three basic factors of individually paced instruction, active response, and immediate feedback. Also, both attempt to create a self-sufficient instructional program which will produce mastery of a subject when used as the sole or primary vehicle of instruction. In addition, since both have been applied primarily to the programming of logically inter-related verbal or symbolic material, both tend to be based on a logically sequenced arrangement of the elements of the subject, in which new concepts are built up gradually on the basis of more elementary material through which the student is first carefully led, in order to try to make sure he masters each prerequisite step before proceeding. (Both can thus be differentiated from the approach mainly followed by Pressey's students and currently exemplified in a number of "self-scoring tests" which seek only to provide adjunctive study-quiz materials rather than a comprehensive, fully ordered and self-sufficient instructional package.) Finally, it may be noted that both Crowder's approach and that of the Skinner school may be implemented more or less advantageously either with more complex mechanical devices for controlled presentation and regulation of the program, or by means of an inexpensive book-format in which some degree of automaticity and control is sacrificed.

"Despite these similarities the differences between the two approaches are very considerable. One tends to use multiple-choice questions, the other stresses 'constructed' responses composed by the student. This difference is fairly consistent, though it should be noted that it is not necessarily an inevitable concomitant of the use of branching versus nonbranching sequences. A second difference is in the fact that the Crowder programs use longer frames (more verbal expository material for each question to which the student responds); this also means that more of the learning from the program is left up to mediation by implicit rather than overt response. The Crowder programs also show a tendency to make many of the steps in the program represented by each frame difficult enough to challenge the superior student, allowing other students to make errors which are then corrected by digression to a remedial sequence. This is a basic difference in rationale which underlies the two approaches. The capability of the Crowder type of program for providing branching sequences is also felt essential by some in order to explain to the student the basis for his errors—errors which the Skinner type of program tries to prevent from occurring in the first place by making all steps easy enough for virtually every student to get them correct.

"A final crucial difference between the two approaches, partly related to the foregoing differences, is the role which is emphasized for the overt responses made by the student as he works through a program. The combination of features that distinguishes most of the branching programs devised by Crowder and those who follow his general approach implies that this role is primarily one of testing to check on the knowledge which the student is expected to have acquired through implicit responding to textual material in each frame of a program. On the contrary, a basic assumption underlying the rationale employed in the programs constructed by Skinner and his associates (as well as numerous others who have begun writing self-instructional programs following his general approach) is that elicitation of appropriate overt responses by the student is *per se* a process through which the responses will be most effectively learned. In this rationale, moreover, the assessment or testing function is minimized, particularly to the extent that the programmer succeeds in achieving a program which can remain motivating to brighter students while preventing even the slow student from making more than a minimal number of errors.

"It may be noted that the emphasis on learning through active response implied here is not restricted to implying a theoretical position which gives primacy to immediate reinforcement as the most directly crucial condition for effective learning to occur. The same emphasis follows also from contiguity theories, such as those enunciated by Guthrie, Estes, and others, which stress only the need to get the appropriate response made by the learner. From either point of view, it is the programmer's aim to elicit the desired responses *in order to effect learning thereby* rather than primarily for the purpose of testing to see if learning has occurred."

The following comments are made concerning research requirements: "To resolve these and other issues in the rationale for program construction and in the determining of optimal device characteristics, research needs to be formulated at the level of dealing with basic variables in the psychology of human learning as these could apply in meeting practical educational objectives. Such research can not only contribute to our basic understanding of the learning process but can also, because of the degree of control afforded in programmed instruction, be translated directly into guidelines which can make direct contributions to the achievement of more effective conditions of learning by students in our schools. . . . "It is unlikely that over-all 'evaluative' comparisons of loosely defined 'types' of devices or classes of programs will contribute much of permanent usefulness to either the science of learning or the technology of programmed instruction. What is needed, rather, is a vigorous effort in programming development, backed up by carefully conceived experimental studies of basic variables (on which future programming procedures and device characteristics can be based)."

Lumsdaine, A. A. "Use of Self-Instructional Devices." *Human Factors Methods for System Design.* (Edited by J. D. Folley.) Pittsburgh, Pa.: American Institute for Research, 1960. p. 291-326. L (c)

This article, appearing in a handbook that presents techniques for the design of human-factor components in complex systems, complements a more general chapter by the same author on "Design of Training Aids and Devices" (1960a). In addition to giving a general review of methods and devices for self-instruction and describing illustrative devices, this report considers specific requirements for self-instructional devices suitable for the training of specific classes of skills, including identification training, procedural-demonstration training, motor skill training, training for variable procedures and diagnostic skills, and automated training for complex job activities.

Lumsdaine, A. A., and Gladstone, A. I. "Overt Practice and Audio-Visual Embellishments." *Learning from Films.* (Edited by Mark A. May and A. A. Lumsdaine.) New Haven, Conn.: Yale University Press, 1958. p. 58-71.

This study was a replication with variations of the earlier study by Hovland, Lumsdaine, and Sheffield (1949), using the same subject matter, but differing in various aspects of experimental procedure and criterion measures employed. The results replicated the previous findings, indicating superiority for active-response procedure over passive review; in addition, the results showed that a simple, straightforward verbal presentation of the subject matter was more effective than one in which various audio-visual embellishments had been used in an attempt to make the presentation more interesting and to provide mnemonic associations.

Maccoby, N., and Sheffield, F. D. "Theory and Experimental Research on the Teaching of Complex Sequential Procedures by Alternate Demonstration and Practice." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication No. 516, 1958. (Library of Congress Card Catalog No. UG633.S96) p. 99-107.

The experiments reported in this paper, which form part of a series conducted at Boston University under contract with AFPTRC, manipulated conditions that affect the probability of a correct response during the learning of sequential procedures. Filmed demonstrations of the procedures that were to be learned (a construction or assembly task) were presented several times in conjunction with practice of the procedure. Three complete demonstration and practice trials were given in all cases. The variation which affected the probability of correct response was the "size-of-step" or length of sequence demonstrated before each practice period. Results showed that the poorest performance, both during practice and on later tests, was obtained when the entire task was demonstrated before it was practiced. In most of the experiments, the best performance was obtained when demonstration of only a short segment of the task was followed at once by practice of just that segment, before the next segment was demonstrated. Results of a related experiment indicated that it was even better to start the students with small demonstration steps, and then transition to larger steps in later trials. For superior students, the best results were obtained with a self-pacing procedure in which the student himself regulated the length of the demonstration segment. Under these conditions, the superior students tended to adopt the transitional pattern spontaneously, i.e., starting with short easy steps on the first trial but ending with whole practice by the last trial. A more detailed report on these experiments and a number of related ones on the role of active student practice in learning is in preparation under the auspices of the National Research Council.

McNeil, J. D. "The Great Didactic and Automated Teaching." Unpublished paper. (Los Angeles: University of California, 1960.) U

In this as yet unpublished paper, McNeil points out some interesting parallels between current approaches to the programming of self-instructional materials and earlier formulations of rules of teaching —particularly those set forth in the seventeenth century by Comenius (John Amos Komensky). Among the rules given by Comenius for effective teaching were the following: "(a) All things should be taught in due succession and only one thing should be taught at a time. (b) Nothing must be learned by rote; comprehension is the light of memory and we should see that whatever has been presented is also understood. (c) Whatever is taught should be taught straightforwardly, and not in a complicated manner. (d) Order, position, and connections of objects should be studied. (e) A subject should not be left until it is thoroughly understood. (f) Stress should be laid on the differences which exist between things, in order that knowledge may be clear and distinct. (g) Learning should proceed from the known to the unknown. (h) Instruction should be fitted to the comprehension of the learner. (i) Children must learn to do by doing. (j) In every subject that consists of several parts, these parts should be linked together as much as possible." McNeil comments on the relationship between these rules and the principles and procedures currently being utilized by programmers and by those who have attempted to formulate provisional rules of programming. [cf. Skinner and Holland, 1959; Gilbert, 1958b; Beck, 1959; among others.]

Mager, R. F. Preliminary Studies in Automated Teaching. *Proceedings of the Institute of Radio Engineers*, October 1959. p. 1790. (a)

In this introductory article to the field of teaching machines Mager cites some of the studies reported in this volume. He describes them as "dramatically effective," predicting that electronics people will be intimately associated with them in both developing and teaching the subject of electronics.—*E. B. Fry et al.*

Mager, R. F. "Teaching: Today and Tomorrow." *IRE Student Quarterly*, September 1959. p. 31-37. (b)

This is a well-written nontechnical account of some principles of learning in relation to automation in teaching. Mager discusses the promise of automation, some of the notions of programming, and the need for research and development in specialized equipment for automated instruction.

Mathieu, G. "Automated Language Instruction: A New Deal for Student and Teacher." *Automated Teaching Bulletin*. 1:5-9; December 1959.

In this article, Dr. Mathieu, Director of the Language Laboratory at Pomona College, identifies and briefly discusses the advantages of "automated language instruction" employing language-laboratory equipment. [See Rheem Califone Corporation *Automated Teaching Bulletin*.]

May, M. A., and Lumsdaine, A. A. *Learning from Films*. New Haven, Conn.: Yale University Press, 1958.

This volume describes experimental studies of factors influencing the effectiveness of instructional films, and thus complements the source references in this field by Hoban and Van Ormer (1950) and by N. E. Miller and collaborators (1957). The papers include experiments dealing with factors of pictorial quality and color, use of live dialogue vs. offstage narration, readability of the instructional commentary, use of overt practice (see separate abstract of paper by Lumsdaine and Gladstone), use of active-response questions in a film for motivational and practice purposes, use of attention-directing instructions, instructional effects as a function of maturity of pupils and amount of previous instruction received on the topic of the film. A basic study of cue and response functions of pictures and words demonstrated that in the learning of paired associates consisting of paired representations of objects, pictures made better stimulus terms but words made better response terms. Some detailed analyses of specific test responses to various verbal-pictorial patterns are also of interest to those concerned with the design of instructional materials in programmed form.

Mayer, S. R., and Westfield, R. L. *A Field Try-Out of a Teaching Machine for the Training of SAGE Operations*. Technical Memorandum, Operational Applications Laboratory, AFCRC, Bedford, Mass., October 1958. (OAL-TM-58-16) △ L

This article describes a field tryout of the Subject Matter Trainer as an on-the-job proficiency-maintaining device, in the SAGE system. Intercept directors and track monitors were trained in knowledge of components critical to development and maintenance of operator proficiency. An alternate group used workbook methods. Both groups made gains; some favored the workbook group. Opinion and attitude survey results were favorable and the authors concluded that such a device could be recommended for use in job knowledge training in field situations.—*E. B. Fry et al*

**Mayhew, D. J., and Johnson, A. F. "Teaching Machines." *Journal of Engineering Education* 50:51-52; 1959. [p. 422-24]

**Mellan, I. "Teaching and Educational Inventions." *Journal of Experimental Education* 4:291-300; March 1936. [p. 265-74]

Melton, A. W. "The Science of Learning and the Technology of Educational Methods." *Harvard Educational Review* 29: 96-106; Spring 1959. (a)

Melton asserts that "the proper question for consideration is not *whether* the science of learning and the management of learning in the classroom can be mutually supporting, but *how* this desirable—even necessary—relationship can be achieved." He states some of the assumptions and inferences underlying this assertion, some difficulties in the way of achieving a mutually supporting relationship, and some ways for overcoming these difficulties. Melton defines the field of learning in broad terms to encompass "all forms of relatively permanent modifications of behavior resulting from experience, with perhaps the exception of those modifications commonly identified as sensory adaptation and fatigue. This means that the science of learning must organize our knowledge and understanding of the acquisition of attitudes, motives, effective and emotional responses, mental sets, simple and complex discriminative acts, serial verbal and motor acts, motor and perceptual skills, meanings, concepts and abstractions, and various cognitive capabilities that go under such names as ideational problem solving, thinking, reasoning, decision-making, and even creative invention." He indicates that the relationship of education to psychology is analogous to the relation between engineering and the physical sciences or medical practice and the biological science. However he points out that "education is not just the straightforward application of the science of learning, any more than engineering is just a straightforward application of the physical sciences. Just as physicists discovered things that made the creation of television possible, but engineers created television, so likewise psychologists have discovered many things about the learning process, but educational technologists must design curricula and teaching machines that exploit that knowledge."

The relationship between basic research and application is a reciprocal one: "Thus, new discoveries in science make new achievements in technology feasible, and recognized sources of deficiency in technology guide the explorations of scientists." Melton discusses the recent relative isolation of educational technology and the science of learning, as evidenced in relatively low overlap in organizational membership and publication channels. He stresses the importance of what Glaser (1960a) has termed a *rapprochement* between educational psychology and the science of learning. Melton feels they have drifted apart partly as a result of "the lack of an explicit strategy for holding them together at the level of advanced training and research."

He draws a parallel between the needs for close interaction in the field of education with a similar interaction between basic science and application which has developed in military psychology. Among the difficulties to be met, he identifies first the immaturity of the science of learning, in which "advances in empirical generalizations and in the theoretical integration of these generalizations have resulted only in *islands* of knowledge and understanding within the science of learning"; second, the inadequacy of our present taxonomy of tasks or skills; and third, confusion concerning the proper functions of applied science or technology, on the one hand, and research on the other hand. This confusion is grounded in two misconceptions—that basic and applied research are somehow antithetical (whereas actually they are complementary), and second, that the kind of learning investigated determines whether research is basic or applied (which he points out is not true).

As requirements for a profitable working relationship between the science of learning and the technology of classroom instruction, he identifies the need for a formal statement of the relationship between the two, and a concerted planning effort in which the objective is to formulate the technological and scientific problems in particular areas of education. Such planning efforts should include identification of the present "state of the art" in the management of particular types of learning, the identifying and ordering in priority of needs for improvement in such management, the identification of which aspects of problems are amenable to basic and applied research, and identification of the needs for systematic development efforts in which the available state of the art can be employed. A third major requirement is adequate facilities and subjects for research and development studies, including ready access to school populations for analytic studies. Finally, he says, "there must be a re-examination of the training of psychologists and educational psychologists with these opportunities in mind." His general conclusion is that ". . . the present low level of interest and effort in the application to the classroom of recent advances in the science of learning is a state of affairs which can be corrected. Furthermore, it seems to me that vigorous effort to apply the necessary corrective measures to both the science of learning and the technology of education is in the best interests of a healthy science of psychology and a healthy profession of education. But over-riding all these considerations is the conviction that the full realization of the human resources of this nation cannot be achieved without a science-based management of learning processes of our children. It seems to me, therefore, that we should get on with the business, so that within a few years we can know that the goals of educational psychology are at least as well supported by educators, psychologists, educational psychologists, and all of our citizens, as are the goals of military psychology."

Melton, A. W. "Some Comments on 'The Impact of Advancing Technology on Methods in Education' by Dr. Simon Ramo." Setember 1959. U (b)

[The following comments were made by Dr. Melton following an address by Dr. Ramo at the Annual Convention of the American Psychological Association, Cincinnati, September 1959. Dr. Melton has indicated that these comments also apply to the 1957 paper of Dr. Ramo, which is reprinted on pages 367-81 of this volume.]

My comments will relate particularly to Dr. Ramo's suggestions regarding the application of engineering technology to education, with specific reference to teaching machines. I am sure that the "brave new world" of automated education described so graphically and hopefully by Dr. Ramo captures the imagination of all of us. This is not to say that all thinking men will be pleased with the prospect of this extension of mechanistic science and engineering to the moulding of minds through the marriage of behavioral science and electronics engineering. Surely, some humanists will see the work of the devil in this effort to automate education, just as they have seen the work of the devil in some applications of nuclear physics. However, the worries of these people who cherish the old order and fear the new will not be mine, nor do I think they should be psychology's. Nevertheless, I can remain calm in the face of this impending educational revolution, and I think that you should remain calm, only if we assume that, in addition to engineering feasibility, there will be two guiding considerations in the development and use of automated aids to human learning of the complex sorts involved in the production of educated men and women. The first of these considerations relates to educational objectives; the second relates to the development of the science and technology of psychology.

First, with respect to the educational objectives, I interpret Dr. Ramo's interests to be the promotion of an educational technology which can be employed as an efficient means toward whatever educational objectives our society defines as "good." When viewed as a technology for achieving a higher level of education in more people in less time and with lesser brainpower cost, if not manpower cost, I see no possible basis for worry. However, I do see a basis for worry if there is failure to recognize that a taxonomy of educational objectives reveals many objectives other than the acquisition of knowledges and skills. I refer, of course, to such objectives as the inculcation of socially acceptable attitudes and motivations; the inculcation of motivation to achieve, to create, to invent, to think independently; and the development of skills in interpersonal relations.

When viewed against this background of educational objectives which literally encompasses the whole spectrum of behaviors associated with the concept of the civilized man, it is clear that the initial successes of automated education will be in the achievement of knowledges and skills. In fact, many mechanical aids to such achievement have already been devised and used extensively. It is now clear that the state of the art in engineering is adequate to support much more complicated types of knowledge and skill acquisition than the simple rote-learning and problem-solving aspects of reading, writing, and arithmetic, and so we are now witnessing attempts to automate language, higher mathematics, the fundamentals of science, history, etc. But, let it not be forgotten that there is the problem of appropriate *utilization* of these knowledges and skills by the individuals who will possess such elegant mental furniture as a gift from electronic engineering. Such individuals must be intellectually curious, inventive, and ingenious; they must be socially and emotionally mature in the employment of their intellectual talents; and above all they must be motivated to perform their roles in our ever more complex society. What I am attempting to say is: Let us *not* consider the goals of this educational technology to be merely, for example, the attainment of mastery of differential calculus in the greatest number of people in the shortest possible time; instead, let us be sure to consider also the techniques for improving the motivation of individuals for the attainment and appropriate use of such knowledge and the techniques for accelerating the acquisition of attitudes and social habits that are just as necessary for the advancement of our society as are those skills.

Some of you will perceive, I am sure, that I am a bit worried lest automated education become a fad, to take its place alongside progressive education, the mental testing movement, and the guidance movement. By putting automated education in such company I mean no offense to it or to progressive education, mental testing, or educational guidance; but I do mean to say that all represent piecemeal approaches to the technology of education, all represent emphases on some educational objectives to the exclusion of other equally valid objectives, and that all represent applications of psychological science that have been—or, in the case of automated education, might be—adopted and lavishly supported by our society while the furtherance of the basic knowledge and theory on which they should be grounded has been forced to scrounge for itself. Which leads me to the plea that—this time, at least—we attempt to anchor this new component of the technology of education in the complete system of educational objectives and in the science of psychology. Let us keep always before us the full range of educational objectives that is defined, in the last analysis, in terms of the desired behavior of the whole person in our society—let there be no exclusive emphasis on the perfection of intellectual skills at the one extreme, *nor* on the encouragement of individual initiative and originality at the other extreme, *but rather* an emphasis on both of these as well as all other desired characteristics of the whole educated man, as revealed by our taxonomy and systematic integration of educational objectives. Also, and with comparable emphases, let us plant this new component of the technology of education (machine aids to learning) firmly in the science of psychology, and use the goals of the technology as justification for the liberal support by our society for advancement of basic knowledge and theory in areas of psychology most necessary for valid decision-making within the technology.

My principal concern has to do with the second of these exhortations; that is, the support of basic psychological science in a manner and to a degree appropriate to the emphasis given automated education. Perhaps I am a "worry wart," but I can see the managers of resources in our society, as well as some starry-eyed educators and psychologists, being carried away by the imaginativeness and insightfulness of the concept of automated education, and forgetting our counsel that contemporary psychology is able to give only first approximations as answers to questions about the processes and programming of human learning which are quite critical in the engineering of education. As a consequence, I can publicly cringe at the thought that millions may well be appropriated for such technological efforts, with the usual 1 percent, or one-tenth of 1 percent, for fundamental research in psychology in support of the improvement of the technology. Why am I gun shy on this point? Those of you who know the history of psychological research in the Air Force personnel and training effort over the past ten years will understand. In the context of the Personnel and Training Systems of the Air Force, it was recognized about 1950 that a technology was needed, *and feasible*, for anticipating the tasks that human beings would perform in new weapon systems, like the ballistic missile systems, and for specifying, far in advance of even the prototype model, what the specific selection, training, and on-the-job management techniques for the production and maintenance of these human components should be. Thus, the technology promised the delivery of the essential human components of a weapon system concurrently in time with the delivery of the equipment components.

After a period of confusion and poor communication, the very simple *idea* of extracting information about human tasks from the *designs* of future weapon systems and using such information to establish the special procedures for selection and training of men to perform those anticipated tasks finally caught on, and was properly recognized as an essential step in the production of an operational weapon-system. Consequently, millions of dollars are now being spent annually by the Air Force in implementing this procedure for defining what is now appropriately called the Personnel Subsystem of each new weapon system. But somewhere along the line a warning of the psychologists who devised this technology was forgotten. Lost somewhere along the line was their warning that the technology that could be supported by contemporary psychological science was primitive—only a first approximation—and that substantial support of certain critical areas of fundamental psychological science should proceed concurrently with the initial implementations of the technology, if the technology were to become reliable and effective. Such support was not only not increased within the Air Force; it was drastically decreased. Now we are witnessing the elaborate employment of a technology by the Air Force, but with the scientific underpinning on starvation rations. One is reminded of the disembodied grin of the Cheshire cat, and the whole affair would be the source of some sardonic humor if the consequences were not so serious for our long-range security, for the viability of the important technology which was introduced into the Air Force, and for our science.

Now, this is not the only case where enthusiasm for a psychotechnology has been coupled with relative apathy for its scientific foundations. I think here again of mental testing, progressive education, guidance, visual education. My concern is that this not be the history of the next ten years in the automation of teaching. This is neither the time nor the place for me to discuss the details of the basic research that should receive support. . . . It is important to note, however, that psychologists and educators are currently joining forces in thinking about the objectives of education and about the developments in basic knowledge and theory of learning that are necessary to support these objectives. These efforts will come to naught if the resources for the development of basic knowledge do not represent some substantial and continuing proportion of the total resources applied to the application of engineering technology to education.

In summary, I believe that the full exploitation of electronic engineering in a radical revision of the technology of education is an essential condition for achieving effective control over the learning process. I am, therefore, a willing partner to Dr. Ramo's enthusiasms and prognostications. At the same time, I think we will be fostering an abortive and potentially harmful fad—and certainly a very costly fad—if we fail to temper our exploitation of gadgetry in terms of clearly formulated educational objectives, only some of which will be promoted by such gadgetry. Which is to say that the educational objective to be achieved should determine the use of such devices, rather than the availability of the device determining, or limiting, the objective. This matching of educational methods and media to the components of an intricate system of objectives is no simple exercise and will tax the talents of psychologists and educators for many years to come. One principle reason why this is so is that the scientific understanding of human behavior, and especially the scientific understanding of the learning, motivation, and performance aspects of human behavior, can provide only first approximations to answers to critical questions involved in the statement of educational objectives and to critical questions involved in the design of a teaching machine.

From this line of thought I derive my second plea—or "warning," if you will—that some substantial proportion of the resources employed in the exploitation of engineering for the benefit of education be allocated to basic and applied research on the learning process, on human motivation, and on related problems of human behavior. If this is not done, the fruits of this imaginative adaptation of engineering technology to the technology of education will be largely wormy and rotten, and only occasionally fit for human consumption. In view of the high economic cost of the automation of education, such occasional pay-off cannot be afforded, and might well result in abandonment of the whole idea. On the other hand, if the definition of educational objectives and the advancement of basic knowledge and theory in psychology can move forward at an accelerating pace to match the pace of engineering applications to education, I am prepared to view the new technology of education which Dr. Ramo has so forcefully described, as a conceptual "break-through" of the first magnitude, and as a unique opportunity for the educator, psychologist, and engineer to work together for a technology that will make the new world *much* braver than Huxley ever imagined.

Meyer, S. R. "A Program in Elementary Arithmetic: Present and Future." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) Chap. 6, p. 83-84. New York: John Wiley & Sons, 1959.

This is an early report of an investigation by Meyer using a slider type machine to teach arithmetic to elementary school children. The machine was built by IBM. . . . Meyer used 26 children attending grades one through six. The study was more of an investigation procedure than a controlled experiment. Meyer found that a fairly large minority of second-grade children had difficulty in matching answers correctly. . . . She also found that the prompts were not always obvious to a child. Sometimes the child must learn the prompts. Meyer was concerned with the problem of understanding; to teach this, she used Skinner's rich context idea of number concepts. This meant that multiplication, division, and addition of simple concepts (lower numbers) were taught before progressing on to higher numbers. She noted that in the school where she was conducting the investigation the children were chanting the counting of fives up to 100 before she had given them the concept of "nine" and the various combinations leading up to the concept "nine."—*E. B. Fry et al.*

**Meyer, S. R. "Report on the Initial Test of a Junior High-School Vocabulary Program." (1960). [p. 229-46]

Michael, D. N., and Maccoby, N. "Factors Influencing Verbal Learning from Films Under Conditions of Audience Participation." *Journal of Experimental Psychology* 46:411-18; December 1953.

This article (an earlier and somewhat more detailed version of which was also reported by Michael, 1951) was a follow-up to the study of Hovland, Lumsdaine, and Sheffield (1949). It had the purpose of investigating some specific factors bearing on the effectiveness of active student response in learning from instructional sequences. The experiment employed a film dealing with subject matter of the substantive verbal (continuous-discourse) type, concerning defense against atomic attack. Active student response was provided at intervals between sections of film exposition by the insertion of question-and-answer sequences. Responses were written (rather than oral as in the case of previous studies of active response —Hovland, Lumsdaine, and Sheffield, 1949; Lumsdaine and Gladstone, 1958). Feedback following student response was given by announcement of the correct answer to each question. The study focused on three aspects of the use of Socratic review sequences: (a) motivational as against practice effects; (b) overt as against covert responding; (c) the importance of feedback or response-correction. The dominant role of the practice factor was shown by comparing the results for the passive or no-participation group with those for active response groups. The results of criterion tests administered following the instruction were analyzed separately for test questions which were directly covered by the practice questions used during instruction and those which were not covered. The results showed that gains from active response were primarily attributable to specific practice effects rather than the motivational benefits accruing to items not specifically practiced. Other comparisons in the experiment lead to the conclusion that the single largest factor in determining the efficacy of the active-response procedure was the provision of feedback-correction through giving the correct answer after each question had been responded to by the students. In this experiment, covert or subvocal response appeared to be about as effective as overt written response.

Miller, N. E., and others. "Graphic Communication and the Crisis in Education." *Audio-Visual Communication Review* 5:1-20; Summer 1957. (special issue)

This issue is a summary of research findings and implications of theory and research for the design of instructional materials, with particular reference to films and related audio-visual media. [See also the summaries by Hoban and Van Ormer, 1950; Cook, 1960; and Allen, 1957.]

Miller, R. B. *Handbook on Training and Training Equipment Design.* Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, 1953. WADC Technical Report 53-136. △

This *Handbook* is intended to aid in preparing recommendations on the design and use of training equipment. As such, it permits cross referencing to a companion report, WADC Technical Report 53-138, *Human Engineering Design Schedule for Training Equipment.* The contents of the *Handbook* include learning and transfer theory, principles applicable to problems of training, and bibliographic references. One principal theme which is developed is that different degrees of learning have different implications for transfer of training and the best form of presenting knowledge of results. Stages of learning are analyzed in detail, as are the variables in "knowledge of results." The principal sections are titled: I. Human Learning—An Overview; II. The Role of the Instructor in Training; III. The Trainer as a Demonstrator of Principles; IV. The Use of Knowledge of Results; V. The Problem of Simulation; VI. The Problem of Motivation; VII. Preparing the Specifications for a Training Device. The many problems indicated as requiring further research in the field of human learning and training suggest the importance of "programmatic" studies. These materials, in conjunction with WADC TR 53-137, *A Method for Man-Machine Task Analysis,* and WADC TR 53-135, *A Method for Determining Human Engineering Design Requirements for Training Equipment,* are designed for use by trained personnel.—*R. B. Miller.*

Miller, R. B. *Psychological Considerations for the Design of Training Equipment.* Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, December 1954. WADC Technical Report No. 54-563. △

This 139-page report was prepared under a Human Engineering Factors Project administered by the psychology branch of the Aero-Medical Laboratory at Wright Air Development Center under Gordon Eckstand. The report covers a wide range of topics—principles of learning and transfer of training, physical simulation, knowledge of results and scoring, proficiency measurements, and an outline of steps in designing training devices.

Of interest is the comparison of learning versus action feedback. Learning feedback is defined as that which modifies the next response after a period of time, while action feedback is defined as an instantaneous feedback which modifies a continuing action. The difference seems to be the use of memory. Miller makes an interesting distinction also between operational and artificial feedback. Artificial feedback is feedback which is found only in training situations, that is, the instructor cautions the student, or a click is heard when on target. Miller holds that the use of artificial cues is not justified. For example, making the target red aided learning, but when the red was removed, the performance dropped back to the level of the control group which had equal practice without the artificial cue. Miller found that natural cues predominated over symbolic cues unless the student was trained otherwise. Learning to disregard misleading cues is also important and training is helpful here. If two cues are equally relevant, learning to use one does not mean that the other one is also learned. Caution should be exercised that spurious cues are not learned. Dramatizing cues such as emotional music played during a training movie may increase motivation and interest, but do not aid learning *per se.* Miller states that a difficulty ratio in training devices should provide for approximately 50 percent success with progress towards the end.—*E. B. Fry et al.*

Morton, F. R. "The Language Laboratory as a Teaching Machine," *Language Teaching Today.* (Edited by F. J. Oinas.) Report of the Language Laboratory Conference held at Indiana University January 22-23, 1960. Publication No. 14 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics. (In Press 1960.)

This 85-page article discusses the psychological principles of language instruction. Skinner's notions on teaching machines are described, and the author points out the teaching-machine aspects of a language laboratory. Interesting experimental studies on language learning are presented, and a plan is detailed for the development of a language-laboratory teaching machine.

Newman, S. E. "Student Versus Instructor Design of Study Method." *Journal of Educational Psychology* 48:328-33; October 1957.

This article, a published version of an earlier, 1956, AFPTRC Maintenance Laboratory Staff Memorandum entitled "The Design of Materials and Procedures for Learning," compared techniques for the learning of names for 20 electrical symbols. Two groups were used, a student group and an instructor group, with 15 subjects in each group. The student group was merely asked to study the materials independently. The instructor group was given a prescribed routine in which the materials were grouped into four steps of five items each. For each set, each of the five items was presented individually, being pronounced first by the instructor and then by the student, and the name of the set (depending on the characteristics with a symbol; e.g., straight lines) was identified. This was followed by a procedure in which the student was asked to name each of the five symbols, being corrected if he erred, and then was tested by having each of the five symbols presented without corrections (one trial each). After this, each of the 20 symbols was presented alone, followed by presentation of the symbol paired with its proper name. With either one or two repetitions of the instructor-group procedure, self-study for an equivalent period of time was found to produce a higher degree of learning. The same was true in a further experiment in which the procedure for the instructor group was revised to give more prompting to the student. The author concludes that: "The effective design of materials and procedures for tasks even so apparently simple as learning the names of electrical symbols is not easy. Much information is already available from laboratory research on learning which might prove useful in designing materials for such tasks. However, until considerable experience is accumulated about how such information can best be applied to the design of materials and procedures for learning in classroom type situations (or until additional research-based information is available), the practice of allowing the student to decide on the techniques he will employ during study of the kind of task used here may be a defensible one.

Pask, G. "Automatic Teaching Techniques." *British Communications and Electronics* 4:210-11; 1957. (a)

This is a brief article describing some of the characteristics of teaching machines in general. More specifically it mentions the Solartron automatic teacher which was developed in England and has been applied to the learning of keyboard machines such as business and adding machines. Such devices are described more fully in a later article by Pask (1958) on pages 336-48 of this book. Also mentioned is another type of automatic teacher which " is concerned with the ability to do mental commercial arithmetic which is largely dependent on being able to see groups of numbers (e.g., in a list of figures to be added together) as a 'whole.' The machine teaches this ability by providing graded exercises consisting of 'sums,' the answers to which can be given in a keyboard form. The 'helpful' information, in this case, is provided by giving partial answers." Pask goes on to say that, "Managerial training situations in certain complicated industrial control operation procedures (e.g., chemical plant control) can also be catered for. These processes resemble games of strategy because the human operator is doing his best to keep charge of the situation; they are well suited to application of the automatic teacher. Radar training and aircraft interception techniques can also be 'learnt,' and other control tasks such as high-speed flying. . . ."

Pask, G. "A Teaching Machine for Radar Training." *Automation Progress* 2: 214ff.; May 1957. (c)

A machine is described for the training of radar operators. Electronic circuits are linked directly to a radar target simulator so as to replace the human instructor. "The Equipment is known as I. D. M. A. (Interactive Decision Making Assembly) and . . . helps to speed up training routines by making the target 'strategy' progressively more difficult, until the pupil makes a mistake, in which case the target interception is made easier to achieve. . . . It is the distinctive feature of I. D. M. A. that input data can be valuated without any hard and fast rules being programmed into the machine. What is, or is not, found difficult by the human operator cannot be predicted and, therefore, cannot be written into a program."

"Where the structure of the system is known, input can be in definite numerical data; where only the function is known, data will be of a probabilistic kind. If the function is only partly known, a feedback system is necessary to overcome the effects of random variables in the environment. For instance, an automatic pilot is based on a definite model of the plane's behavior, with feedback to allow for variations in external conditions. Where no model of the system exists, a device like I.D.M.A. must be employed to replace the human observer. It starts with a 'blank page' and constructs a continuously changing model, as it goes along. Construction and adaptation of such a model is a completely unconstrained learning process, the only determinant for decisions made being the purpose to be achieved. As implied in the term 'interactive,' a trial-and-observation feedback is involved. In a teaching machine, this feedback corresponds to the training routine, and is deliberately maximized. Decisions made by such equipment are not 'true' in the sense in which those based on a definite model can be true. They are, however, useful, because they lead to efficient control of the observed system."

In order to simplify description of the radar training system, target motion is limited to one geometrical plane, neglecting the height coordinate, considering only a single target. "The radar simulator supplies two A.C. potentials, each representing one of the coordinates describing the position of the operator's 'plane' as it is represented in a display. The teaching machine then selects a target position for the next instant which is computed, *relative* to the operator's position, and resolved into angular coordinates, appropriate to the radar display, before it is fed back into the simulator."

**Pask, G. "Electronic Keyboard Teaching Machines." *Education and Commerce* 24:16-26; July 1958. [p. 336-48] (a)

Pask, G. "Teaching Machines." Unpublished paper, Solartron Electronic Group, Limited, 1958. (Mimeo.) U (b)

This is a paper presented at a symposium at the Second Congress, International for Cybernetics in Namur, Belgium, in September 1958, which describes, in general, two main topics discussed in the 1958 and 1960 papers by Pask in this volume [p. 336-48, and 349-66]. They are (a) a description of the basic notions and design behind keyboard devices of the "SAKI" type, and (b) a brief discussion of the game-interaction model between the trainee and the teaching machine.

Pask, G. "The Self-Organizing Teacher." ***Automated Teaching Bulletin*** 1 (No. 2): 13-18; December 1959. (a) [See Rheem Califone Corporation]

A teaching machine must have some similarity to the man that it teaches; that is, it must be able to carry on a conversation or to interact. In a practical sense, this means a man interacting with a self-organizing computer. Pask feels that this man-machine interaction will achieve a stability that neither one would obtain alone. In this system of two self-organizing systems, one at least can be accurately analyzed (the machine) and thus new insights be gained in the learning phenomena. Hence a self-organizing computer which is allowed to build its own mode of an efficient tutor, based on the man's learning responses, should prove an effective device for training and for understanding learning.—*E. B. Fry et al.*

Pask, G. "The Teaching Machine." *The Overseas Engineer*, February 1959. p. 231-32 (b)

This is a brief article which provides a good description of the Solartron Automatic Keyboard Instructor ("SAKI"), which is a teaching machine that continually examines the trainee's performance, assesses it, and looks out for defects. "The way in which it accomplishes this feat depends upon the skill being taught. In the case of punching business-machine cards, for example, the skill is to translate lines of numerical or alphabetical exercise material into manual movements and key depressions executed [See Pask (1958a) p. 336-48]. Performance is measured in terms of rate and accuracy, by comparing in a coincidence circuit the sequence of keys which are depressed, with the ideal sequence of keys which should have been depressed indicated electrically by a programmed template. The template comparison determines whether for each item a move of any kind has been made and, if it has, whether the move was a correct response or an error response. A rate measurement is made separately for each item by comparing the interval which is allowed for making a move determined by the teaching machine, and corresponding to the interval when the visual indicator remains at the item, with the interval actually needed to make a move. Supposing the move is a correct response, the difference between these intervals is measured and represented as a potential which will be defined as $Sji(t)$, meaning the measure obtained at the i-th position on the j-th exercise line at an instant, t. If the move had been an error response no rate measurement would have been made and the value of $Sji(t)$ would have been defined arbitrarily as minus one.

"Modifying the Training Routine. The teaching machine must react to the results obtained in terms of achieving its objective, by making different modifications of the training routine. In 'SAKI' there are four alternative exercise lines each with 24 items. These are scanned in a regular sequence so that the device is unable to select a particular exercise line. It may, however, determine the number of repetitions of an exercise line, and thus the number of occasions on which the trainee must deal with this line before he is provided with the next line in the sequence. There is also an averaging circuit which derives the quantity θ, defined as the average over many items of the quantities $Sji(t)$. The objective is achieved if no value of $Sji(t)$ is low and if the value of the quantity θ is high.

"The teaching machine may thus determine: (a) The time interval during which the indicator remains at any item in the exercise line being scanned. This necessarily determines the interval allowed for dealing with this item. (b) The amount of helpful information which is provided; that is the extent to which the helpful information display lamps are delayed. (c) The number of repetitions of each exercise line before the next line in the sequence is selected. (d) The average rate at which the exercise line as a whole, rather than its component items, must be dealt with.

"Clearly, the modifications (a) and (b) must depend upon quantities which are determined for each position in each exercise line, and these can be defined as a set of variables $Ajt(t)$ of which there are 4 X 24. The values of these variables are registered on a set of 4 X 24 storage condensers. When the indicator scans across an exercise line the scanning mechanism also connects an amplifier input to one of these storage condensers for each item. This amplifier provides a sequence of values; that at the i-th position on the j-th exercise line being a Ajt the modifications of (c) and (d), on the other hand, depend upon a measure of performance averaged over a set of items. This measure has already been defined as θ and is not registered separately for each item.

"It is possible to show that although the activity of the trainee upon his own is unpredictable and although the activity of the teaching machine upon its own is either trivial or unpredictable, a union of the two, interacting to form a combined system, is a measurable entity. In particular, as the teaching machine competes with the trainee and makes him accept a more difficult training routine and deal with it successfully the learning process is depicted in terms of the distribution of high valued entries in $A(t)$. These may be recorded and the pattern in the recording used to predict further states of the combined system."

**Pask, G. "Adaptive Teaching with Adaptive Machines." (1960) [p. 349-66]

Peterson, J. C. "A New Device for Teaching, Testing, and Research Learning." *Transactions of the Kansas Academy of Science* 33:41-47; 1930.

This reference was taken from the Angell 1948 study. It describes two procedures for giving immediate knowledge of results. The first device was an envelope which contained several layers of cardboard. When a pin was punched through several selector holes it was possible to record the response, and knowledge of results was given to the testee depending on whether or not the pin went through the envelope. This is believed to be the forerunner of Pressey's punchboard.

The second device is a small answer card which has five spots of chemical after each number. The testee responds to a multiple-choice question by marking a special ink on top of the spot. If the spot is the incorrect answer the ink shows a red mark; if it is the correct answer, the spot turns blue or a definite dark color. Pressey, in personal correspondence, later referred to this as the Chemo card. . . . The stylus used in marking these Chemo cards is a cheap fountain pen filled with a special non-expensive ink. Pressey seemed rather interested in the Chemo card and suggested that "apparently a number of fairly well-known schemes of secret ink may so serve, and your chemistry department might help you out."—*E. B. Fry, et al.*

**Peterson, J. C. "The Value of Guidance in Reading for Information." *Transactions of the Kansas Academy of Science* 34: 291-96; 1931. [p. 52-58.]

Popp, H., and Porter, D. "Programming Verbal Skills for Primary Grades." Paper read at the annual convention of the Department of Audio-Visual Instruction, NEA, Cincinnati, March 1960. *Audio-Visual Communication Review* (in press).

This paper describes and illustrates programming techniques used for pre-reading training with young children and for the teaching of spelling and reading in primary grades.

**Porter, D. "A Critical Review of a Portion of the Literature on Teaching Devices." *Harvard Educational Review* 27: 126-47; 1957. [p. 114-32]

**Porter, D. "A Report on Instruction Devices in Foreign Language Teaching." Prepared for the International Training and Research Program of the Ford Foundation, October 1, 1958. (Mimeo.) [p. 186-205] (a)

**Porter, D. "Teaching Machines." *Harvard Graduate School of Education Association Bulletin* 3: 1-5; March 1958. [p. 206-14] (c)

Porter, D. "Some Effects of Year-Long Teaching Machine Instruction." *Automatic Teaching: The State of the Art* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 7, p. 85-90.

In the academic year 1957-58 an experiment was carried out on the teaching of spelling to elementary school children under the sponsorship of the U.S. Office of Education. Porter used a simple write-in type of teaching machine. The materials were printed on eight-inch-wide duplicating paper. On the first line of the sheet a question was printed which required the student to respond directly on the sheet through an open window in the machine. The student then moved a lever which rolled the paper up one line; this placed the question and the student's written answer under a sheet of glass, also exposing the correct answer so that he could score himself right or wrong. The lever was then moved to bring the next item into view. The chief functions of the machine were to expose one item at a time and to prevent cheating.

An interesting aspect of the entire program was that no spoken instruction was given to students. Teaching materials paralleled standard lessons in textbooks which were used by the control group as closely as possible. The words taught were exactly the same as those received by the control group.

The teaching-machine program attempted the following sequence of responses:

1. Identification of spelling words in a meaningful sense context.
2. Matching the correct word or given definition.
3. Matching several words on the basis of letter structure.
4. Writing missing letters in words presented in sentence context.

Twenty-two weeks of a normal 34-week program of spelling instruction were given on both the sixth- and the second-grade levels using the teaching machine. Results were measured on standardized achievement test scores. These showed the experimental group to be significantly superior.

Porter reported that there was essentially no relationship between intelligence scores and achievement in the experimental group, but there was a significant positive relationship in the control group. Porter's experiment seemed to indicate that this type of instruction is particularly good for the lower IQ groups. No relationships existed in either group with respect to sex or whether or not students liked the instructional method. To check on the "novelty" factor of the new instruction process, Porter compared first-half and second-half machine scores; no differences were apparent.

In the sixth-grade experimental group a positive relationship existed between the number of responses per lesson required of subjects and subsequent achievement, but no relationship existed between the number of errors per lesson and achievement. In the sixth-grade group, lesson materials were available on a week-by-week basis; yet the experimental group spent only about one-fourth as much time studying as did the control group.

In the second-grade group, materials were available on an *ad libitum* basis for 12 weeks. During this time the spread of materials extended to nine lessons; in other words, some students had a high rate of response when released from the more strait-laced type of curriculum. . . . The experimental group gained 1.42 years while the control group gained .90 years. A sign test showed this to be significant at the .01 level. Statistical data for the second grade showed that the experimental group gained .80 years while the control group gained 1.10 years. (U test, .005). A second method of rating second-graders was letter position scores in which each letter was scored. In this the experimental group gained 65.6 while the control group gained 37.8 (U test, .001). In an interesting sidelight Porter found that 20 percent of the second-graders and 10 percent of the sixth-graders could not match their answers with the answers given in the machine.—*E. B. Fry et al.*

**Pressey, S. L. "A Simple Apparatus Which Gives Tests and Scores —and Teaches." *School and Society* 23: 373-76; March 20, 1926. [p. 35-41]

**Pressey, S. L. "A Machine for Automatic Teaching of Drill Material." *School and Society* 25: 549-52; May 7, 1927. [p. 42-46.]

**Pressey, S. L. "A Third and Fourth Contribution Toward the Coming 'Industrial Revolution' in Education." *School and Society* 36: 668-72; November 19, 1932. [p. 47-51]

Pressey, S. L. "Further Attempts To Develop a 'Mechanical Teacher.'" Paper presented at the annual meeting of the American Psychological Association, Philadelphia, September 1946. U [Abstract in *The American Psychologist* 1: (No. 7); July 1946]

This is a review of the work of Pressey and his students as of 1946.

**Pressey, S. L. "Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-Instruction." *Journal of Psychology* 29: 417-47; April 1950. [p. 69-88]

Pressey S. L. "Some History Regarding Self-Instructional Devices." Paper read at Symposium on the Automation of Teaching, the American Psychological Association Convention, Washington, D.C. August 1958.

This was a general historical review of the work of Pressey and his students [see also papers by Briggs (1949), Jensen, Jones, Little, Peterson, Severin, and Stephens]. Pressey drew three major conclusions from his experience: First, to be educationally effective, the use of a self-instructional device must become an integrated, habitual part of the student's work. Second, to achieve such wide use, devices should be as simple and convenient to use as possible. Third, the more complex devices are, in Pressey's view, primarily useful in situations in which such features as cumulative measures of progress or selective review are required.

Pressey, S. L. "Certain Major Psycho-Educational Issues Appearing in the Conference on Teaching Machines." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 16, p. 187-98.

Pressey stresses the fact that much relevant research has been done in educational psychology that is often overlooked by teaching-machine researchers. He is also concerned with the problem of using teaching machines as an "aid to" or a "replacement of" regular instruction.

Pressey disagrees with what he calls the "short-easy-step-program type of thinking" exemplified by Skinner and his followers. He cites such instances as the Ford Foundation report, *They Went to College Early*, which showed that bright students who skipped up to two years of high school did better work than most college students. He concluded with a plea for more experimentation with control groups and practical applications.—*E. B. Fry, et al.*

**Pressey, S. L. "Some Perspectives and Major Problems Regarding 'Teaching Machines.'" (1960) [p. 497-505]

Price, G. R. "The Teaching Machine." *Think* 25: 10-14; March 1959.

In a popular introductory article, Price discusses Skinner's principle of "shaping" and the teaching machines used by Porter and Holland at Harvard. Price calls them "an important new invention."—*E. B. Fry et al.*

**Ramo, S. "A New Technique of Education." *Engineering and Science Monthly* 21: 17-22; October 1957 [p. 367-81]

Ramo, S. "A New Technique of Education." *Institute of Radio Engineers Transactions on Education* E-1: 37-42; 1958.

(This is a reprinting of Ramo's 1957 *Engineering and Science Monthly* article.)

Ramo, S. "The Impact of Advancing Technology on Methods in Education." Paper read at the Annual Convention of the American Psychological Association, Cincinnati, September 1959.

This paper gave a further exposition of some of the ideas presented in Ramo's 1957 paper [p. 367-81], with particular emphasis on the requirements posed, and tools potentially provided for education, through "the extending of the human intellect by electronics." The need for examination of objectives and careful analysis of ways of meeting them was pointed out as prerequisite for effective automation of functions. For a comment on Ramo's ideas, see Melton, 1959b, and Skinner, 1959a in this Appendix.

Rath, G. J.; Anderson, N. S.; and Brainerd, R. C. "The IBM Research Center Teaching Machine Project." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 11, p. 117-30.

This paper is primarily a discussion of binary arithmetic; the most interesting section, from a teaching-machine standpoint, is a description of the use of an IBM 650 general purpose digital computer to simulate a teaching machine. This computer used a typewriter input and output. The subject matter was the teaching of binary arithmetic. Psychological principles programmed into the teaching-machine computer were:

1. The student was required to construct his own answers by typing them into the machine.
2. The machine corrects the answer unit by unit, or digit by digit.
3. Errors are eliminated in some cases, as when the student tries to make the answer too long. Here the machine says "Right" before an error can be typed.
4. The machine allows for individual differences in skill level and rate. In other words, if the student is making a large number of errors, the machine will go back and repeat steps. The number of errors made determine how far back the machine will skip back for repetition.

Knowledge of results is provided automatically as the student continues working through the program. If the student makes an error, the machine stops. Knowledge of results is given, character by character. The computer also allows the student to skip problems if no errors are made. The skipping of problems is optional with the student. The IBM 650 has the capacity to handle a number of input stations of this type. A cost advantage is seen in using a 650 computer for research purposes, since it is much cheaper to program the characteristics of a teaching machine into a computer than to build a similar device as a teaching machine. If it is anticipated that the characteristics are likely to be changed or should be changed as part of a research project, perhaps it would be actually cheaper to use the 650 for research. But as a regular training device the cost of the IBM 650 would be prohibitive.—*E. B. Fry, et al.*

Reid, L. S. "A Review of *Automatic Teaching: The State of the Art,* E. H. Galanter, Editor." *Science* 131: 29-30; 1960.

This is a brief review of Galanter's book, which concludes that, "The book's most important contribution is in the reiteration of the vast number of questions that must be answered before such techniques can even approximate their maximal usefulness, and the emphasis on the role teaching machines can play as a research tool in providing the answers."

Renner, Inc. *Inexpensive Missile Training—A New Approach.* Paper read at the U. S. Navy Guided Missile Training Symposium, Naval Ordnance Laboratory, Whiteoak, Md., March 1958.

In this paper Renner describes and discusses training possibilities of the "study card" set, also known as the "trainer-tester." Van Valkenburgh, Nooger, and Neville hold the basic patent. Renner claims that the study card set provides realistic and accurate synthetic simulation of problems and situations encountered in the operation and maintenance of real equipment, a technique referred to as paper simulation. Features of the study card set are well-thought-out problems; step-by-step testing, as well as training; immediate reinforcement; active learning; and self-scoring.—*E. B. Fry, et al.*

*Rheem Califone Corporation. *Automated Teaching Bulletin* 2; December 1959.

This publication, distributed by the Rheem Manufacturing Company, is described as a quarterly journal devoted to basic and applied research in automated teaching and training. Two issues have appeared, September 1959 and December 1959. For abstracts of the papers in the latter issue, see Mathieu, Fry, Pask, Corrigan, Schutz, Silberman and Coulson, and Freeman. The December issue also contains brief notes on a number of current research projects concerned with automated teaching and related areas. All of the contents of the first (September) issue are reprinted in the second issue.

Rigney, J. W.; Bryan, G. L.; and Svenson, D. W. "The Multiple Choice Item as a Training Device." Paper read at the annual convention of the American Psychological Association, Chicago, September 1956. U

This paper was a preliminary report on the study reported by Bryan, Rigney, and Van Horn, 1957.

**Rothkopf, E. Z. "Heuristic Discussions of Psychological Bases for the Conduct of Training by Automatic Devices: I. A Functional Diagram and an Approach to Research." Unpublished Staff Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo. April 1957 [p. 305-10] (b)

**Rothkopf, E. Z. "Some Research Problems in the Design of Materials and Devices for Automated Teaching." [p. 318-28] Paper read at Symposium on the Automation of Teaching, American Psychological Association Convention, Washington, D. C., August 1958. (b)

Rothkopf, E. Z. "The State of an Art?" (A review of *Automatic Teaching: The State of the Art*, edited by E. H. Galanter.) *Contemporary Psychology* 5: 104-105; 1960.

This is a critical review which emphasizes the tentative character of many of the papers included in the book, and questions the contents of the book as furnishing an adequate representation of the state of the art. The reviewer feels that "The papers which deal with descriptions and illustrations of various instructional techniques are the most consistently substantial and most generally useful of the book."

Saltzman, I. J. "Comments on Teaching Machines." *Language Teaching Today.* (Edited by F. J. Oinas.) Report of the Language Laboratory Conference held at Indiana University, January 22-23, 1960. Publication No. 14 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics. (In press, 1960)

This is a commentary on the paper by Skinner given at the same symposium. Saltzman emphasizes the changing role of the teacher as a result of automation in instructional procedures.

Schutz, R. E. "Progress and Preparation of Automated Instructional Material and Criterion Measures." *Automated Teaching Bulletin* 1, No. 2: 31-33; December 1959. [See Rheem Califone Corporation.] (Also presented at Symposium on Automated Teaching, Western Psychological Association Convention, San Diego, Calif., April 1959.)

Schutz reports the results of approximately one year of an elementary arithmetic program developed according to "Skinnerian principles." Although no direct rules on programming are reported, it is stated that the investigator can successively approximate a refined instructional sequence, (that is, learning by doing). Usual paper and pencil and tests are not considered applicable for automated instruction programming because "we do not have direct knowledge of the student's work through the program."—*E. B. Fry, et al.*

Severin, D. G. "Appraisal of Special Tests and Procedures Used with Self-Scoring Instructional Testing Devices." *Abstracts of Doctoral Dissertations* 66: 323-30; 1955. Columbus: The Ohio State University Press.

The self-instructional device used was the "punchboard" previously described (in the 1950 paper by Pressey). The experimentation was in classes totaling about 2000 students in The Ohio State University and Otterbein College. Subject matter included Russian and hard English vocabulary, naval terms, a short scientific article, and contents of a course in psychology. Instructional tests were 4, 3, and 2 choice, application, judgment; end-tests were always different from the instructional tests they followed. For example, Russian practice tests were multiple-choice, but the end tests recall or write-in.

The punchboard increased the range in number of errors and so presumably the discrimination of less from more able students; thus 10-90 percentile range in a psychology class on a 4-choice 30-item test was 32 with the punchboard but 8 when the test was simply marked as in usual testing—as might be expected, since by this last procedure only one error per item can be made, but three can be when the punchboard is used. In short, simply as a testing device, the punchboard not only eliminates scoring but improves measurement. In addition, all end-tests with all types of subject matter showed the punchboard to have brought about some learning—and more, in proportion as the subject matter was meaningful. The punchboard was proven very easy to learn to use: after only five minutes instruction in its use, a class made practically the same median and distribution of scores, on both practice test and end-test, as another class with three months' habituation in use of the device.

In both Russian and English vocabulary, 4-choice instructional tests yielded a greater range of error-scores but also took more time than 2-choice tests; however, end-test means and distributions were practically the same. If learning rather than testing is the major purpose, only 2-choice instruction tests thus seem needed—with saving in both test-taking and test-making time. End-tests showed less learning with a 2-choice instructional vocabulary test than with a "two-pair" test in which each line had a Russian word at beginning and end and their English meanings between them, the student to designate (by punching the first or second hole in the corresponding line on the punchboard) which meaning went with the last Russian word. Each line of the two-pair test thus contained two right associations, and each Russian word appeared twice in the test. Efforts to devise special types of test for self-instructional purposes seem clearly warranted.

One class studied a list of 30 naval terms for 15 minutes, then spent 15 minutes on a practice test covering these terms using the punchboard; a second class spent the full 30 minutes simply studying the list; a third class studied the list for 15 minutes and then for 15 minutes studied a copy of the practice test on which the correct answers had been marked—the practice test was thus seen but not reacted to. All three classes then took a recall end-test, and another end-test a week later. An experiment of similar design used a short scientific article as the material to be learned. Results were in favor of the punchboard more often than not, but differences were all negligible. It was inferred that for short relatively easy tasks, automation of learning might be of little value—though a score useful in other ways might thus easily be got.

At least with meaningful matter, once through a practice test with punchboard yielded as good end-test results a week later as a second time through either the same day or next day; presumably the important cues are got the first time through. Practice tests should deal with a topic as it is to be used: if English meanings of Russian words are to be learned, a Russian-to-English practice test is more effective than English-to-Russian. It is concluded that (a) such a device as the punchboard, usable for class experiments, can greatly facilitate educational research, and (b) distinctive useful instructional tests and procedures may thus be developed.—S. L. *Pressey.*

Silberman, H. F. "Introductory Description of Testing Machines (Physical Characteristics)." Paper read at Symposium on Automated Teaching, Western Psychological Association Convention, San Diego, Calif., 1959. U

Silberman states that automated teaching devices consist of four units: (a) an input unit, (b) an output unit, (c) a storage unit, and (d) a control unit. These devices differ in complexity from digital computers to paper-and-pencil devices. The goals vary from rote memory to complex decision-making skills. Inputs would include audio and visual inputs, verbal inputs and motor inputs. Outputs include audio and visual display, and knowledge of results in the form of information, food, coins, and other reinforcers. Stored information concerning the behavioral data differs from machine to machine. Machines also differ in type of problem material, accessibility of materials, and instantaneous storage capacity. Control units may determine the selection of materials as a complex function of the student's performance. At the simplest level, presentation is made merely of a series of pre-stored items. Performance may be evaluated as right or wrong, or in more complex ways.—*E. B. Fry et al.*

Silberman, H. F. "The Automation of Teaching." *Behavioral Science.* 1960 (In press).

The article discusses the desirability of using computers with teaching machines. One of the important values mentioned is that the computer can be programmed to modify its own instructions in response to student performance. An interesting and complex diagram on "Functional Steps in Problem Selection Program" shows the many factors and choices a computer can make before selecting a given item for the student. [A paper by this title was also presented by Silberman at the September 1959 Convention of the American Psychological Association.]—*E. B. Fry.*

Silberman, H. F., and Coulson, J. "A Draft Summary of Findings in an Exploratory Teaching Machine Study." *Automated Teaching Bulletin* 1: 35-37; December 1959.

This is essentially an abstract of the study reported by Coulson and Silberman on p. 452-68 of this volume.

**Skinner, B. F. "The Science of Learning and the Art of Teaching." *Harvard Educational Review* 24: 86-97; Spring 1954. [p. 99-113]

Skinner, B. F. *Report #1 on the Development of Methods of Preparing Materials for Teaching Machines.* Unpublished Report to HumRRO, January 1957. △ U (b)

This brief report is a summary of some of the problems considered and some of the solutions proposed in Skinner's early work on the development of materials for a constructed-response type of teaching machine. The first problem to be considered is: What to teach. In beginning to develop programmed material, some indication of the student's adequacy should be obtained. "A suggested solution is a review disk, or set of disks, as a refresher for all students, with 'branching' or further work for those who have difficulty. It is also possible to present a whole series of disks to all students, assuming that review is not aversive to the already trained student. This latter method has the advantage of preventing a fairly well-prepared student from continuing with a hidden weakness in some small part of the skill."

A second problem discussed by Skinner concerns the body of materials to be taught. In this connection he notes that: "By using many sources of materials, we can program a larger number of the descriptive and explanatory statements on a series of discs than any one textbook author is likely to present within the textbook limitations of space and capacity to sustain student interest."

The third topic considered is the framing of questions. Skinner says that: "Successive frames must generally be independent of each other for two reasons: (a) no referral to previous material is possible while the machine is in operation, and (b) those questions which have been answered correctly twice are not presented again. Material such as figures, formal definitions, or examples can be presented on a panel which is always available to the student. . . . Another suggestion for introducing texts is a scrambled reference book containing hints (perhaps for more than one subject matter) to which the student is referred by a number exposed in a central section of the question slot on the first round only."

Several kinds of frames are suggested by which the programmer is able to control the presence or absence of textual prompts. One kind that is suggested is the use of "Two-form" frames. These are frames, in which, for example, the definition is given in two ways with a response blank in one of the ways in which the definition is presented; or perhaps a definition plus an incomplete example or application of this definition; or perhaps a completed application or example plus an incomplete one. Another type of frame suggested is one in which prompting is not on the basis of prior intraverbal connections, but on the basis of new ones acquired in immediately preceding frames.

Skinner, B. F. *Report #2 on the Development of Methods of Preparing Materials for Teaching Machines*. Unpublished Report to HumRRO, April 1957. △ U (c)

This report is a brief discussion of some notions involved in generating and ordering the frames in a program. "Probably an ideal set of rules for programming should possess at least the following characteristics: the principles should be *general* enough to apply to any subject matter; they should be 'mechanical,' or exact enough so that there is no 'art' involved; the rules should be as simple as possible and thus *communicable;* and finally, of course, the program should be *effective*.

"When one wishes to transmit knowledge about a topic, he makes a number of statements about it. For example, if one wishes to teach about acceleration, he makes statements such as, 'Acceleration is the time rate of change of velocity', '$v=vo+at$ is a linear function if a is a constant,' etc. When a programmer has at hand a number of statements such as these, he constructs (by coding) variations on them. The statements which result may be classified into two kinds: 'event series' and 'fading series.' *Event Series*. This kind of a series results when we isolate one statement about an event and ask the question 'In how many ways can this statement be made?' An example will illustrate the consequences: suppose we wish to teach the relationship, $a=(v-vo)/t$. We can see immediately that a large number of equivalent but different statements can be generated: $at=v-vo$; $t=(v-vo)/a$; $v=at+vo$; $v-at=vo$; acceleration is the time rate of change of velocity; the product of acceleration and time elapsed is equal to the change in velocity; time elapsed equals the change in velocity divided by the acceleration; . . . etc. We have, then, a large number of statements which are equivalent in the sense that they all say the same thing, but which differ in the symbols used or the order in which the symbols appear. We will refer to a series of statements such as this as an *event series* to differentiate them from a *fading series*." The notion of fading or vanishing is also defined here [See Skinner, 1958 c, pages 138-58].

Considering the ordering of cycles the report talks about event cycles and fading cycles. The learner may be presented with several cycles. The advantages of this method of programming are stated to be the following: (a) the student will get an over-all view of whatever the cycle covers; (b) in a fading series, items that vary in difficulty or importance can be dropped out on succeeding cycles. This can help in changing context for the statements which remain; (c) the difficulty of a program may be manipulated by the judicious elimination or addition of cycles without altering its content.

Skinner, B. F. "Final Report to Human Resources Research Office." Unpublished report, October 1957. △ U (d)

Skinner reports that a large part of the work for the year, under contract with HumRRO, consisted of testing various ways of presenting and programming material. This is necessary in order that programming may be converted from an art to something approaching a science. "In brief, what is needed is nothing less than an analysis of various forms of knowledge. We must decide upon the behavior we want to impart to the student. This is, in itself, an innovation. It does no good to say that we are interested in setting up mathematical 'ability' or artistic 'appreciation' or a language 'skill.' We must identify the relevant behaviors much more closely if we are to set the goal of the program in useful terms. Once the material to be programmed has been defined, it must be broken down into an optimal number of steps in an optimal order."

Skinner suggests several ways in which supplementary material can be used with a program:

"(a) Material can be supplied which is to be read first, but is not present during the operation of the machine. (b) Material can be available the first time a set of frames is presented but not again. Such material can show hints, references to lists of rules, references to manuals or rule books, and so on. (c) Other forms of material can be present during the operation of the machine. Maps, texts, diagrams, and so on, can be supplied with each set of frames; also, auditory material has been used in teaching dictation. The student receives a disk for this machine and another for a selective photograph. A sector on the machine may instruct the student to play passage C-7. The student listens to this passage as often as he likes. When he has decided he knows what it is saying, he writes it down. He then operates the machine and the text of what the passage is actually saying is revealed. He may continue to listen to the passage while reading the correct response."

An interesting discussion of notion of "leading" in a program is contained in the following quotation:

"The student is first asked to talk about familiar things—for example, dots, lines, distances, directions, angles, and so on—using his everyday vocabulary. He is led—much as a lawyer 'leads' a witness—to discuss relations among these, many of which he has probably not previously noticed. Technical terms are then slowly inserted. For example, in geometry, the student first discusses dots; a dot then comes to indicate merely a point in space and the discussion proceeds in terms of points. Within a short time, and with little effort, a student finds himself discussing many of the materials of geometry or trigonometry in technical terms. . . . His behavior may begin to 'pay off' in the discovery and demonstration of unexpected properties or relations. . . . What he discovers is inherent in the material he talks about. The leading process is familiar to any good teacher who has had the opportunity of closely following individual students or small homogeneous classes. . . .

"An example of how the machine can carry the student beyond a merely mechanical production of sentences may be taken from the field of Physics. The equation

$$s=\frac{1}{2}gt^2$$

can be learned like '54°40′ or fight'—say, by the vanishing technique; but we want the student to be able to do more than reproduce the formula. Yet what more do we want? If we want him to say what *s*, *g*, and *t* stand for and to explain the numerical operations indicated by ½ and ^{2} we can build up appropriate verbal responses. When this has been done, we can say that the student 'knows what the terms in the equation mean.' We might do this simply by teaching the sentence, 'The distance a body falls from rest is equal to one-half the value of the gravitational constant times the square of the time in which the body falls,' as we might teach a perfectly arbitrary line of poetry. Surely something more is wanted. One thing is the interpretation of examples. The student must be able to see the relevance of the formula or the symbols, or the prose statement of the law, even when none of the terms or symbols appears in the description of the example. . . . The student must be taught to pass from instance to definition and from definition to instance either by setting up explicit verbal behavior case by case or appealing to generalization from a few instances. We might still ask something more—namely, that the student see that this law is similar to other laws, or that the terms appearing in it have dimensions similar to those in other laws, and so on. . . .

"If these examples suggest that we are neglecting some well-known goal such as the inculcation of 'abilities' or 'understanding' or 'insight,' then those who defend these goals should be asked to specify them in terms of the behavior which they would accept as indicating a successful realization. If, among these specifications there are kinds of behavior which have not been mentioned here, the process of programming knowledge will need to be extended."

Skinner, B. F. "Programming Verbal Knowledge for Automated Instruction in Schools and Colleges." Paper read at Symposium on the Automation of Teaching, American Psychological Association Convention, Washington, D. C., August 1958. U (a)

This paper was a preliminary presentation of some of the material on programming presented in Skinner's 1958 (c) article in *Science*.

**Skinner, B. F. "Teaching Machines." *Science* 128: 969-77; October 24, 1958. [p. 137-38] (c)

Skinner, B. F. "Comments on 'The Impact of Advancing Technology on Methods in Education' by Dr. Simon Ramo." September 1959. (a)

In comments made following the address by Dr. Ramo, at the Annual Convention of the American Psychological Association, Cincinnati, September 1959, Skinner made the point that the engineer should avoid duplicating with equipment the standard conditions in education. For example, he need not consider that he is to teach a class, there being no longer any reason to teach more than one student at once. Devices are needed to allow the individual to proceed at his own rate. Skinner indicated that the incorporation of automatic means of branching at this time seem unnecessarily expensive, as compared with a simple system whereby, after an hour's work on the machine, the next set of materials given to the student, depending upon what he has done, can be shifted between sessions.

In his comments Skinner also took issue with the statement by Gilbert [p. 475-85], that the first thing to do in designing a teaching machine is to go and watch a teacher. He indicated that, rather, what should be analyzed as a basis for designing a teaching program is the student and the contingencies needed to bring about changes in his behavior.

Skinner, B. F. "The Programming of Verbal Knowledge." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 4, p. 63-68. (b)

Skinner believes that the shaping of a written response through "successive approximation" is inefficient. Successive approximation here refers to the manner in which he teaches pigeons to play ping-pong. That is, a subject is rewarded each time he closely approaches the goal; the subject is then rewarded the next time only if he gets closer to the goal, and so on in ascending approximation. Skinner does not consider this effective in shaping verbal responses. He prefers the technique known as "vanishing." In vanishing, the total word is first presented; then one small part is removed for the succeeding item. On the second item, two small parts are removed, and so on, until the whole word is finally removed, or "vanished."

In addition to vanishing, Skinner observes that sometimes a student may combine known elements for a response. Moreover, the use of a term or fact in another syntactical (sentence order) arrangement makes a good type of item. An important goal is the enrichment of "understanding." Skinner here defines understanding as the use of a fact under several different conditions, which is a process of recombining. Hence, just answering a question correctly is not necessarily the goal.

The method of juxtaposition is used in combining a nonverbal stimulus, such as a symbol, with a word. A verbal response may be brought under verbal control by the method of "equating," an example of which would be "X means Y," or use of the word "or," or use of substitution. He believes that much of learning a new language is of this nature.—*E. B. Fry et al.*

Skinner, B. F. "Special Problems in Programming Language Instruction for Teaching Machines." *Language Teaching Today.* (Edited by F. J. Oinas.) A report of the Language Laboratory Conference held at Indiana University, January 22-23, 1960. Publication No. 14 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics. (In press, 1960.)

This is a statement primarily concerned with the relationship between Skinner's description of verbal behavior, as presented in his book on this subject, and the programming of materials for language instruction.

**Skinner, B. F. and Holland, J. G. "The Use of Teaching Machines in College Instruction." Final Report to The Fund for the Advancement of Education, August 15, 1958. [p. 159-72.]

Smith, D. E. P. *Speculations: Characteristics of Successful Programs and Programmers. Automatic Teaching: The State of The Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. p. 99-102.

Among numerous comments and examples of programming, Smith outlines the following steps in programming: (a) define the behavior desired, (b) determine the steps to achieve that goal, (c) introduce concepts, (d) induce a web of learning (this refers to the making of many associations in integrating new concepts, similar to what Skinner means by "understanding"), (e) summation (this refers to integrating only items as a whole), (f) review. Review should be done both immediately (every five or 10 frames) and delayed (every 20 to 40 frames).

Smith also suggests that the steps be small, . . . While bright students could skip steps, dull students need more items. He further suggested that more items might not hurt the bright students who can work them very rapidly. He noted that dull students develop "retention and attention" problems after failing several steps. Smith is concerned about the qualifications of a good programmer. . . . He believes that a programmer tends to be analytical, deductive, methodical, perfectionistic.—*E. B. Fry et al.*

Spence, K. W. "The Relation of Learning Theory to the Technology of Education." *Harvard Educational Review* 29: 84-95; Spring 1959.

Spence declares that the time is ripe for a combined research effort between psychologists and educational psychologists, particularly as far as teaching machines are concerned. He believes that programming should properly be undertaken jointly by a subject-matter specialist and a psychologist. He states that teaching machines use immediate reinforcement which is approved by all theorists. The teacher's new role is to build up the student's interest and goals which would serve as secondary reinforcers and incentive motivators. In this respect school rooms are deficient today; teaching machines may make for better reinforcement.—*E. B. Fry et al.*

**Stephens, A. L. "Certain Special Factors Involved in the Law of Effect." *Abstracts of Doctoral Dissertations.* (Columbus; The Ohio State University Press.) No. 64, 1953. [p. 89-93]

Stolurow, L. M. "Teaching Machines and Special Education." An unpublished report of work sponsored by the U.S. Office of Education Cooperative Research Project performed at the University of Illinois. 1960. U

This paper presents a general survey of the development of teaching machines. It summarizes some preliminary findings of research studies, especially as they relate to problems in special education, and describes a current research program on automation in special education. Stolurow describes briefly eight important properties of teaching machines: (a) The machine presents information; (b) the machine requires a response that is overt and measurable; (c) it compares, or permits the learner to compare, the response with a predetermined correct response (comparator function); (d) it provides knowledge of results, and has the potential for doing this after every response (feedback function); (e) it programs information according to an explicit set of rules; (f) it allows for variability in the information-presentation rate (pacing function); (g) it combines learning and measurement into a single set of coordinated operations (collator function); (h) it may automatically reject correctly-responded-to items, so that from trial to trial the learner sees a smaller number (selection function). Not all of these functions are equally important, nor does every so-called teaching machine have all these features. The main variation in these functions to be found among true teaching machines will be in the collator and feedback functions.

Considering implications for research, the author points out that the fact that the teaching machine leaves no ambiguity about the response made by the learner permits the careful investigation and discovery of the differing rate of learning between retarded and normal learners. Other researchable problems that are pointed out are the investigation of reinforcement, the providing of knowledge of results, and teaching-machine pacing. In general, the teaching machine is considered as a device which can help provide an increased understanding of the education process.

Types of teaching machines are classified as follows: (a) nonadaptive machines, such as the memory drum, (b) partially adaptive machines such as the Subject Matter Trainer (cf. Briggs, 1958), and Pressey's punchboard, and (c) adaptive teaching machines which perform the function of branching. In discussing research results, the author's concern with special education and mentally defective students leads him to examine research on mental age and rate of learning, the size of step in a program, and also the type of response required. These things seem to have important implications for education of the mentally retarded. A brief description is given of a program of research currently underway.

Twyford, L. C.; O'Hara, J. G.; and Goldman, A. *Green Light Rater*. Port Washington, N. Y.: U.S. Naval Training Device Center, December 23, 1958. 5 p. Technical Report NTDC 20-0S-3-1. △

(This report is abstracted by Twyford in the Summer 1959 issue of *Audio-Visual Communication Review*.) The purpose of this research and development project was to design and evaluate a teach-test, or self-rating device which would incorporate good psycho-educational principles and which would permit the user to prepare and change the questions.

The apparatus displays 42 cards containing multiple-choice questions. One of four buttons beneath each card glows green if correctly pressed; the others glow red to indicate errors. This rater was found to have the advantage of similar teach-test devices. Production costs were estimated to be below those of similar devices; it was easy to operate and maintain.

Underwood, B. J. "Verbal Learning in the Educative Processes." *Harvard Educational Review* 29: 107-17; Spring 1959.

Underwood points out that many of the so-called facts, laws, or phenomena of verbal learning were derived from studies closely approximating classroom conditions, particularly in the earlier years of educational and learning research. He notes that "group experiments, done in a classroom setting, are not at all unheard of today and for many purposes seem to be quite satisfactory." Underwood feels that factors in verbal learning which account for important differences in acquisition rates would include "meaningfulness, intratask similarity, intertask similarity, active recitation versus passive study, and perhaps one or two others," and would not include such variables as "the effective characteristics of the material, massed versus distributed practice, whole versus part learning, most motivational variables, and many others." He points out that much contemporary work in human verbal learning is not closely oriented to classical theoretical variables, although this was not always the case—e.g., the numerous experiments by E. L. Thorndike. The role of frequency and meaningfulness is discussed at some length, and curves are presented relating association values on nonsense syllables to their occurrence as "natural" three-letter sequences in the English language.

Underwood summarizes his paper as follows: "The role which frequency plays in verbal learning has been discussed. It has been discussed as an *illustration* of how laboratory studies in verbal learning might contribute to an understanding of learning in the schools. It is quite obvious that the present writer believes that the importance of frequency *per se* has been underestimated. But this is not the critical point of the present paper. The critical point is that it does seem possible to analyze the implication of a laboratory variable and to show how it might be relevant to the school situation. It would probably be possible to carry out similar analyses of other variables which have been shown to be of considerable importance for verbal learning in the laboratory. Analyses such as the one made here could do no worse than indicate a few representative experiments to be done in the classroom to determine the degree to which the laboratory principles could be generalized. Such an empirical tie seems nearly mandatory for sound but efficient application of laboratory principles to the real-life situation."

**Weimer, P. K. "A proposed 'Automatic' Teaching Device." *Institute of Radio Engineers Transactions on Education* E-1: 51-53; June 1, 1958. [p. 311-17]

Zeaman, D. "Comments on the Automation of Teaching." Presented at Symposium on the Automation of Teaching at the American Psychological Association Convention, Washington, D. C., August 1958.

In discussing the papers presented by the other members of the symposium, Zeaman drew on his extensive experience at the University of Connecticut in experimenting with teaching machines to teach elementary skills to retarded children. He also discussed use of simple paper-and-pencil devices similar to those of Pressey, developed as self-teaching aids for this and related applications. He raised some interesting questions concerning the relation between theories of learning and the design of self-teaching materials, and presented some data concerning the desirability of presenting the student with varying numbers of alternative answers to the series of questions posed by self-teaching materials. [cf. Lumsdaine, 1958]

Zeaman, D. "Skinner's Theory of Teaching Machines." *Automatic Teaching: The State of the Art*. (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 14, p. 167-76.

. . . Skinner's theory says that the appropriate reference class of experiments, the appropriate experimental paradigm, for construction and programming teaching machines is that of *free operant* conditioning. Others, in particular those using the memory drum, are asserted to be inappropriate. It is the counterthesis of this talk that the existing learning machines (and those likely soon to be built) represent complex admixtures of at least three basic experimental paradigms, *free operant*, *controlled operant*, and *classical conditioning*, and that of these, the free operant may be the least pertinent. Furthermore, of all the commonly used apparati for experimental study of learning, that one which most closely resembles one of Skinner's teaching machines (the write-in model) is the *memory drum*.

By *controlled operant* is meant those experiments incorporating the following features on acquisition trials: (a) Trial begins with presentation of stimulus S_1. (b) One (or in some instances more) incorrect responses are allowed and recorded or not depending upon the particular subclass of controlled operant procedures used. (c) One and only one correct response Rc occurs per trial because Rc terminates S_1 and (d) Rc brings on S_2, a reinforcing stimulus. (e) S_2 does not elicit Rc. (f) Intertrial interval (S_2—S_1 duration) is usually under control of the experimenter (as opposed to the S_1—S_2 interval which is under the subject's control). (g) Acquisition is measured by latency of S_1—*Rc*, or Rc probability (the relative frequency of Rc in time *t*, given S_1 for a series of trials or subjects). Illustrative apparatuses for controlled operant conditioning include the runway, shuttle-box, jumping stand, simple T-maze, discrimination box, and retractible-bar-Skinner-box.

In the *free operant*, similar features are present with this important exception: S_1 is of prolonged duration allowing multiple Rc's (and consequent S_2's) to permit a *rate* measure to be taken. In the *classical* case, an important distinction is of course in the nature of S_2. It is not contingent upon Rc, but rather elicits it. Other features resemble those of controlled operant with this exception: in classical conditioning more than one Rc can occur on a single trial, and in this property it resembles the free operant.

Two of Skinner's machines are analyzed in these terms, the Arithmachine and the Write-in model. It will be found that . . . the Arithmachine encorporates *all* of the features of the controlled operant, but not all of those of the free operant. The Write-in model has elements of classical conditioning, in that S_2 elicits Rc, elements of the controlled operant in that the procedure is trialwise with S_1 terminated after S_2, but no elements of the free operant that do not overlap with those of controlled operant or classical conditioning. It is concluded that there is a large literature of controlled operant and classical conditioning that is relevant for construction and programming of teaching machines, and that Skinner's theory leads to a neglect of this potentially valuable source of pertinent data.—*D. Zeaman*

APPENDIX II

BIBLIOGRAPHY

A

Adams, H. *The Education of Henry Adams.* New York: The Modern Library, 1931. (First published in 1918.)

Adams, H. *The Degradation of the Democratic Dogma.* New York: Capricorn Books, 1958. (First printed by Macmillan in 1919.)

Adams, J. J. "Does the Tachistoscope Help?" *California Teachers Association Journal* 50:20; March 1954.

*Adkins, D. C. "Measurement in Relation to the Educational Process." *Educational and Psychological Measurement* 18: 221-40; Summer 1958.

Allen, W. H. "Readability of Instructional Film Commentary." *Journal of Applied Psychology* 36: 164-68; June 1952.

*Allen, W. H. "Research on Film Use: Student Participation." *Audio-Visual Communication Review* 5: 423-50; Spring 1957.

Almon, T. F. "A Course in Reading Training." *College English* 13: 106-108; November 1951.

Ammons, R. B. *Knowledge of Performance: Survey of Literature, Some Possible Applications, and Suggested Experimentation.* Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, 1954. WADC Technical Report 54-14. △

*Amsel, A. "Some Comments on the Manipulation of Error Responses and Reinforcement Schedules in Self-Instructional Training Devices." Unpublished Staff Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., September 1957. △ U

**Amsel, A. "Error Responses and Reinforcement Schedules in Self-Instructional Devices." (1960) [p. 506-16]

*Angell, D., and Lumsdaine, A. A. *Prompted Plus Unprompted Trials Versus Prompted Trials Alone in Paired-Associate Learning.* Pittsburgh: American Institute for Research, 1960 (in press). △ L

*Angell, G. W. "The Effect of Immediate Knowledge of Quiz Results on Final Examination Scores in Freshman Chemistry." *Journal of Educational Research* 42: 391-94; January 1949.

**Angell, G. W., and Troyer, M. E. "A New Self-Scoring Test Device for Improving Instruction." *School and Society* 67: 84-85; January 31, 1948. [p. 66-68]

Armed Services Technical Information Agency (ASTIA), Document Center. *Title Announcement Bulletin.* Dayton, Ohio: The Center (Knott Building). △

*Arnoult, M. D. *A Comparison of Training Methods in the Recognition of Spacial Patterns.* Lackland Air Force Base, Tex. Air Force Personnel and Training Research Center, February 1956. Research Report AFPTRC-TN-56-27. △

Ash, P., and Carlton, B. J. *The Value of Note-Taking During Film Learning.* (Pennsylvania State College) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, November 1951. Technical Report SDC 269-7-21. △

Ash, P., and Jaspen, N. *The Effects and Interaction of Rate of Development, Repetition, Participation, and Room Illumination on Learning from a Rear-Projected Film.* (The Pennsylvania State University) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, October 1953. Technical Report SDC 269-7-39. △

Atkinson, S. B. *Radio in Secondary Education.* Master's Thesis. Cambridge: Harvard University, 1927. (Typewritten)

Auble, D., and Mech, E. V. "Partial Verbal Reinforcement Related to Distributed Practice in a Classroom Situation." *Journal of Psychology* 36: 165-86; July 1953. (a)

Auble, D., and Mech, E. V. "Quantitative Studies of Verbal Reinforcement in Classroom Situations: I. Differential Reinforcement Related to the Frequency of Error and Correct Responses." *Journal of Psychology* 35: 307-12; April 1953. (b)

Azrin, N. H. "Some Effects of Two Intermittent Schedules of Immediate and Non-Immediate Punishment." *Journal of Psychology* 42: 3-21; 1956.

B

Bamford, H. E., Jr. *The Use of Training Aids in Conceptual Training.* Pittsburgh, Pa.: American Institute for Research, July 1955. △ L

Barbe, W. B. "The Effectiveness of Work in Remedial Reading at the College Level." *Journal of Educational Psychology* 43: 229-37; April 1952.

**Barlow, J. A. *Earlham College Student Self-Instructional Project, First Quarterly Report.* Report submitted to the U. S. Office of Education by Earlham College, Richmond, Ind., 1959. (Mimeo.) [p. 416-21]

*Barlow, J. A. "Project Tutor." *Psychological Reports* 6: 15-20; 1960. (a)

*Barlow, J. A. A review of *Automatic Teaching: The State of the Art,* (Edited by E. H. Galanter). *Harvard Educational Review,* 1960 (in press). (b)

Barry, R. F., and Smith, P. E. "An Experiment in Ninth-Grade Reading Improvement." *Journal of Educational Psychology* 45: 407-14; November 1954.

*Beck, J. "On Some Methods of Programming." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 3, p. 55-62.

Bennett, G. K. *Development of a Morse Code Actuated Printer.* (Psychological Corporation) Washington, D. C.: Applied Psychology Journal, OSRD: 1945. OSRD Report 6233 (PBL 15808). △ L

Berg, I. A. "The Use of Human Subjects in Psychological Research." *American Psychologist* 9: 108-11; March 1954.

Berger, E. J. "Devices for a Mathematics Laboratory." *Mathematics Teacher* 43: 287-90; October 1950.

Berger, E. J. "A Mechanical Device for Drawing the Sine Curve." *Mathematics Teacher* 46: 210-13; March 1953.

*Besnard, G. G.; Briggs, L. J.; Mursch, G. A.; and Walker, E. S. *Development of the Subject-Matter Trainer.* Technical Memorandum, Armament Systems Personnel Research Laboratory, AFPTRC, Lowry Air Force Base, Colo., March 1955. (ASPRL-TM-55-7) △ L

*Besnard, G. G.; Briggs, L. J.; and Walker, E. S. *The Improved Subject-Matter Trainer.* Technical Memorandum, Armament Systems Personnel Research Laboratory, AFPTRC, Lowry Air Force Base, Colo., April 1955. (ASPRL-TM-55-11) △ L

Bijou, S. W. "A Systematic Approach to an Experimental Analysis of Young Children." *Child Development* 26: 161-68; September 1955.

Bijou, S. W. "Methodology for an Experimental Analysis of Child Behavior." *Psychological Reports* 3: 243-50; 1957.

Bijou, S. W. "A Child Study Laboratory on Wheels." *Child Development* 29: 425-27; September 1958. (a)

Bijou, S. W. "Operant or Instrumental Studies with Children Using Baseline Measures." Paper read at the annual convention of the American Psychological Association, Washington, D. C., September 1958. (b)

Bloom, B. S., and Krathwohl, D. R. *Taxonomy of Educational Objectives.* New York: Longmans, Green & Co., 1956.

**Blyth, J. W. "Teaching Machines and Human Beings." (1960) [p. 401-15]

Bortolazzo, J. L. *An Evaluation of Instructional Film Usage in U. S. Navy Training Activities in World War II.* Doctor's thesis. Cambridge: Harvard University, 1949. (Typewritten)

Brandwein, P. F. *The Gifted Student as Future Scientist.* New York: Harcourt, Brace & Company, 1955.

*Briggs, L. J. "Intensive Classes for Superior Students." *Journal of Educational Psychology* 38: 207-15; April 1947.

*Briggs, L. J. "The Development and Appraisal of Special Procedures for Superior Students and an Analysis of the Effects of Knowledge of Results." *Abstracts of Doctoral Dissertations* 58: 41-49; 1949. (Columbus, The Ohio State University Press.)

*Briggs, L. J. *A Troubleshooting Trainer for the E-4 Fire Control System.* Technical Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo. May 1956. (ML-TM-56-16) △ L (a)

*Briggs, L. J. *A Troubleshooting Trainer for the E-4 Fire Control System.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, July 1956. Development Report, AFPTRC-TN-56-94 (ASTIA Document No. 098870). △ (b)

Briggs, L. J. *Design of Maintenance Training Equipment for Fighter-Interceptor Fire Control Systems.* Technical Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., August 1957. (ML-TM-57-16) △ L

**Briggs, L. J. "Two Self-Instructional Devices." *Psychological Reports* 4: 671-76; December 1958. [p. 299-304]

*Briggs, L. J. "Applied Research Problems." Paper read at Symposium on Automated Teaching, Western Psychological Association Convention, San Diego, Calif., April 1959. U (a)

**Briggs, L. J. "Problems in Simulation and Programming in the Design of Complex-Skill Trainers." Paper read at Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Convention, Cincinnati, September 1959. [p. 329-35] (b)

*Briggs, L. J. "Teaching Machines, Education, and Job Skills." *Psychological Reports* 5: 210; 1959. (c)

*Briggs, L. J. "Teaching Machines for Training of Military Personnel in Maintenance of Electronic Equipment." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 12, p. 131-46. (d)

*Briggs, L. J. *A Survey of Cueing Methods in Education and in Automated Programs.* Pittsburgh: American Institute for Research; Research Report AIR-314-60-IR-106, May 1960. (AFOSR-TN-60-286) △ L (a)

*Briggs, L. J. "Teaching Machines." *USAF-NRC Symposium on Education and Training Media.* (Edited by G. Finch.) Washington, D. C.: National Academy of Sciences—National Research Council, 1960. (In press) (b)

Briggs, L. J., and Besnard, G. G. *The Benefits from Increased Practice in Training Electronic System Mechanics.* Technical Memorandum, Armament Systems Personnel Research Laboratory, AFPTRC, October 1955. (ASPRL-TM-55-12) △ L

*Briggs, L. J., and Besnard, G. G. "Experimental Procedures for Increasing Reinforced Practice in Training Air Force Mechanics for an Electronic System." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. Library of Congress Card Catalog No. 56-60063. p. 48-58.

*Briggs, L. J., and DuVall, W. E. *Design of Two Fire Control System Maintenance Training Devices.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, September 1957. Technical Report AFPTRC-TR-57-7 (ASTIA Document No. 134242). △

*Briggs, L. J.; Plashinski, D.; and Jones, D. L. "Self-Pacing Versus Automatic Pacing of Practice on the Subject-Matter Trainer." Unpublished Laboratory Note, AFPTRC, ASPRL-LN-55-8. Lowry Air Force Base, Colo., September 1955. (Mimeo.) △ U

Brown, C. B. "Teaching Spelling with a Tachistoscope." *English Journal* 40: 104-105; 1951.

*Brown, G. I., and Hodgkinson, H. L. "A Note Concerning 'An Application of Recent Developments in Psychology to the Teaching of German' by C. B. Ferster and S. M. Sapon." *Harvard Educational Review* 28: 156-57; 1958.

Brown, J. L. "Vocabulary via Tachistoscope." *Educational Screen* 30: 274, 287; September 1951.

*Bryan, G. L.; Bond, N. A., Jr.; LaPorte, H. R., Jr.; and Summers, S. A. *The AUTOMASTS: An Automatically-Recording Test of Electronics Troubleshooting.* Technical Report No. 11, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, August 1957. (Mimeo.) △ L

*Bryan, G. L., and Rigney, J. W. *An Evaluation of a Method for Shipboard Training in Operations Knowledge.* Tech. Report No. 18; prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, September 1956. △ L

*Bryan, G. L.; Rigney, J. W.; and Van Horn, C. *An Evaluation of Three Types of Information for Supplementing Knowledge of Results in a Training Technique.* Technical Report No. 19, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, April 1957. △ L

*Bryan, G. L., and Schuster, D. H. *An Experimental Comparison of Troubleshooting Training Techniques.* Technical Report No. 30, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, December 1959. △ L

Buchwald, A. M. "Extinction After Acquisition Under Different Verbal Reinforcement Combinations." *Journal of Experimental Psychology* 57: 43-48; January 1959.

Bugelski, B. R. *The Psychology of Learning.* New York: Henry Holt and Co., 1956.

*Buitenkand, N. *Methods of Instruction in Component Location for Troubleshooting in Complicated Equipment and Means for Testing Accuracy of Answers.* U. S. Patent No. 2,764,821. Filed August 25, 1953 (assigned to Van Valkenburgh, Nooger, and Nelville, Inc., New York), patented October 2, 1956.

Burns, John L. *Communications and Education.* Reprint of a speech delivered at the meeting of the National School Boards Association, 1958. Camden, N. J.: Radio Corporation of America, 1958.

Buswell, G. T. *Remedial Reading at the College and Adult Levels: An Experimental Study.* Supplementary Educational Monograph, No. 50. Chicago: University of Chicago Press, 1939.

C

California State Department of Education. *Educational Programs for Gifted Pupils.* Sacramento: the Department, 1960.

*Cantor, J. H., and Brown, J. S. *An Evaluation of the Trainer-Tester and Punchboard Tutor as Electronics Troubleshooting Training Aids.* (George Peabody College) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, October 1956. Technical Report NTDC-1257-2-1. △ L

Carney, C. D. "An Experiment in Teaching 'Mechanical Arithmetic' in an Elementary School." *Business Education World* 32: 128-29; November 1951.

*Carpenter, C. R., and others. *The Classroom Communicator.* (The Pennsylvania State University) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, October 1950. Technical Report SDC 269-7-14. △ L

Carpenter, C. R., and others. *An Investigation of Closed Circuit Television for Teaching University Courses.* (The Pennsylvania State University) State College, Pa.: the University, 1955.

*Carpenter, F. "How Will Automated Teaching Affect Education?" *The University of Michigan School of Education Bulletin*, October 1959.

*Carpenter, F. "The Teaching Machine and Its Education Significance." Paper read at the Twelfth Annual Meeting of the American Association of Colleges for Teacher Education, Chicago, Ill., February 1960.

Carr, A. R. "Using a Calculator to Develop Basic Understandings and Meanings in Arithmetic." *Mathematics Teacher* 43: 195-96; May 1950.

*Carr, W. J. "Bibliography on Teaching Machines." Unpublished report. Lewisburg, Pa.: Bucknell University, 1958. U

**Carr, W. J. *Self-Instructional Devices: A Review of Current Concepts.* Wright-Patterson Air Force Base, Ohio: Aero-Space Medical Laboratory, Wright Air Development Center, August 1959. WADC Technical Report 59-503. [Article appears in this volume under title "A Functional Analysis of Self-Instructional Devices." p. 540-62] △

Carroll, J. B., and Sapon, S. M. "Cross Validation of Foreign Language Aptitude Tests with a Second Air Force Trial Course Sample." *Report No. 2. The Harvard Language Aptitude Project.* Cambridge: Harvard University, February 1955. (Mimeo.) L

Carroll, J. B.; Sapon, S. M.; and Richards, S. E. "Construction and Validation of a Test Battery for Predicting Success in Spoken Foreign Language Courses." *Report No. 1. The Harvard Language Aptitude Project.* Cambridge: Harvard University, June 1954. (Mimeo.) L

Carter, H. L. "A Combined Oral Reading and Psychogalvanic Response Technique for Investigating Certain Reading Abilities of College Students." *Journal of Applied Psychology* 34: 267-69; August 1950.

Cason, E. B. *Mechanical Methods for Increasing the Speed of Reading.* Teachers College Contributions to Education, No. 878. New York: Teachers College, Columbia University, 1943.

Chalmers, E. L., Jr.; Morrison, E. J.; Briggs, L. J.; and Pens, E. H. *Evaluation of a Method for Teaching Troubleshooting Techniques.* Technical Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., December 1957. (ML-TM-57-34) △ L

Chapman, R. L., and Kennedy, J. L. "The Background and Implications of the Systems Research Laboratory Studies." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. (Library of Congress Card Catalog No. 56-60063)

*Cook, D. A. "On Vanishing Stimuli in Instructional Material." *Journal of the Experimental Analysis of Behavior* 3: 1960 (in press).

*Cook, J. O. "Processes Underlying the Learning of a Single Paired-Associate Item." Unpublished Staff Memorandum, Maintenance Laboratory, AFPTRC, Chanute Air Force Base Field Extension, September 1956. △ U

*Cook, J. O. "Supplementary Report: Processes Underlying Learning a Single Paired-Associate Item." *Journal of Experimental Psychology* 56: 455; November 1958.

*Cook, J. O. "Research in Audio-Visual Communication." *Research, Principles, and Practices in Visual Communication.* (Edited by J. Ball and F. C. Byrnes.) East Lansing, Mich.: National Project in Agricultural Communications, Michigan State University, 1960. p. 91-106.

*Cook, J. O., and Kendler, T. S. "A Theoretical Model to Explain Some Paired-Associate Learning Data." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. (Library of Congress Card Catalog No. 56-60063.) p 90-98.

Cook, J. O., and Spitzer, M. E. "Supplementary Report: The Effects of Between-Term Delay and Overt Practice of the Response Term Upon Paired-Associate Learning." Unpublished paper. 1959. U

*Cook, J. O., and Spitzer, M. E. "Supplementary Report: Prompting Versus Confirmation in Paired-Associate Learning." *Journal of Experimental Psychology* 59: 275-76; 1960.

*Cornell, F. G.; Damrin, D. E.; Saupe, J. L.; and Crowder, N. A. *Proficiency of Q-24 Radar Mechanics: III. The Tab Test—A Group Test of Troubleshooting Proficiency.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, November 1954. Research Bulletin, AFPTRC-TR-54-52. △

*Corrigan, R. E. "Automated Teaching Methods." *Automated Teaching Bulletin* 1: 23-30; December 1959.

*Coulson, J. E. *An Experimental Teaching Machine for Research at SDC.* Technical Memorandum, System Development Corporation, Santa Monica, Calif., October 1959. (a)

*Coulson, J. E. "Programming for Individual Differences." Paper read at Symposium on Automated Teaching, Western Psychological Association Convention, San Diego, Calif., April 1959. U (b)

*Coulson, J. E., and Silberman, H. F. "Proposal for Extension of Automated Teaching Projects." Unpublished Field Note, System Development Corporation, Santa Monica, Calif., July 1959. U (a)

**Coulson, J. E., and Silberman, H. F. *Results of Initial Experiment in Automated Teaching.* System Development Corporation, Santa Monica, Calif., July 1959. [p. 452-68] L (b)

*Coulson, J. E., and Silberman, H. F. "Effects of Three Variables in a Teaching Machine." *Journal of Educational Psychology* 51: 1960. (in press)

Counts, G. S. *Education and American Civilization.* New York: Teachers College, Columbia University, 1952.

Cronbach, L. J. *Educational Psychology.* New York: Harcourt, Brace & Company, 1954.

*Cross, K. P., and Gaier, E. L. "Technique in Problem-Solving as a Predictor of Educational Achievement." *Journal of Educational Psychology* 46: 194-206; April 1955.

*Crowder, N. A. "The Concept of Automatic Tutoring." Unpublished Staff Memorandum, Armament Systems Personnel Research Laboratory, AFPTRC, Lowry Air Force Base, Colo., October 1955. △ U

Crowder, N. A. "Self-Instructional Materials for GETS Trainer." Experimental Manual. Unpublished document, *Maintenance Laboratory, AFPTRC,* Lowry Air Force Base, Colo., October 1956. △ U

*Crowder, N. A. *A Part-Task Trainer for Troubleshooting.* Lackland Air Force Base, Tex.: Air Force Personnel Training and Research Center, June 1957. Development Report AFPTRC-TN-57-71 (ASTIA Document No. 131423). △

Crowder, N. A. *An Automatic Tutoring Book on Number Systems*—Vols. I, II, and III. Arlington, Va.: Psychological Research Associates, 1958. PRA Report No. 58-17. △ L (a)

Crowder, N. A. *An Automatic Tutoring Book on Number Systems.* Vol. I. Timonium, Md.: Hoover Electronics Co., 1958. (b)

*Crowder, N. A. "Intrinsically Programmed Materials for Teaching Complex Skills and Concepts." Paper read at Symposium on Automation of Teaching, American Psychological Association Convention, Washington, D. C., August 1958. U (c)

*Crowder, N. A. "Automatic Tutoring by Means of Intrinsic Programming." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 10, p. 109-16.

*Crowder, N. A. *The Arithmetic of Computers.* Published by Western Design, a Division of U. S. Industries, Inc., Santa Barbara Airport, Goleta, Calif., 1960. (a)

**Crowder, N. A. "Automatic Tutoring by Intrinsic Programming." (1960) [p. 286-98] (b)

*Crowder, N. A. "Automatic Tutoring by Means of Intrinsic Programming." *Proceedings of the Educational Testing Service Invitational Conference,* October 1959, on "The Impact of Testing on the Educational Process." Princeton, N. J.: Educational Testing Service, 1960 (in press). (c)

Curtis, F. D., and Woods, G. G. "A Study of the Relative Teaching Values of Four Common Practices in Correcting Examination Papers." *School Review* 37: 615-23; October 1929.

D

Dallenbach, K. M. "The Effect of Practice Upon Visual Apprehension in School Children." *Journal of Educational Psychology* 5: 321-44, 369-404; 1914.

DALLENBACH, K. M. "The Effect of Practice upon Visual Apprehension in the Feeble-minded." *Journal of Educational Psychology* 10: 61-62; February 1919.

*DAMRIN, D. E., and SAUPE, J. L. *Proficiency of Q-24 Radar Mechanics: IV. An Analysis of Checking Responses in Troubleshooting on Tab Test Problems.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center. November 1954. Research Bulletin AFPTRC-TR-54-53. △

*DARBY, C. L. *An Annotated Bibliography on the Automation of Instruction.* Research Memorandum. U. S. Army Air Defense Human Research Unit, Fort Bliss, Tex., July 1959 U. (Mimeo.) △ L

DAVID, E. E., JR. "Voice-Actuated Machines: Problems and Possibilities." *Bell Laboratories Record* 35: 281-86; August 1957.

DAVIDSON, A. O. *Deterrents to Film Use in Schools.* Doctor's thesis. Cambridge: Harvard University, 1952. (Typewritten)

*DAY, J. H. "Teaching Machines." Unpublished paper. (Athens: The Ohio University, 1959.) U

DEARBORN, W. F., and ANDERSON, I. H. "A New Method for Teaching Phrasing and for Increasing the Size of Reading Fixations." *Psychological Record* 1: 459-75; December 1937.

DEARBORN, W. F.; ANDERSON, I. H.; and BREWSTER, J. R. "Controlled Reading by Means of a Motion Picture Technique." *Psychological Record* 2: 219-27; May 1938.

DEESE, J. *The Psychology of Learning.* New York: McGraw-Hill Book Co., 1952.

DE MANDACH, A. B. "Conversational Robotapes." *Modern Language Journal* 38: 138-41; March 1954.

*DIEDERICH, P. B. "Self-Correcting Homework in English." *Proceedings of the Educational Testing Service Invitational Conference,* October 1959, on "The Impact of Testing on the Educational Process." Princeton, N. J.: Educational Testing Service, 1960 (in press).

DODSON, W. L. "Mechanical Aids in a Practical Reading Program." *Texas Outlook* 23: 13; December 1939.

*DOWELL, E. C. "An Evaluation of Trainer-Testers." Headquarters, Technical Training Air Force, Keesler Air Force Base, Miss., June 1955. (Report No. 54-28) △ L

E

EARNEST, E. "Must the TV Technicians Take Over?" *AAUP Bulletin* 44: 582-88; September 1958.

EDWARDS, A. L. *Statistical Analysis for Students in Psychology and Education.* New York: Rinehart & Co., 1946.

EDWARDS, A. L. *Statistical Methods for the Behavioral Sciences.* New York: Rinehart & Co., 1954.

*EDWARDS, W. D. "Skinner's Teaching Machines." Unpublished Laboratory Note, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., May 1956. (Mimeo.) △ U (a)

*EDWARDS, W. D. "The Use of Statistical Decision Functions in Making Practical Decisions." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. (Library of Congress Card Catalog No. 56-60063) p. 115-24. (b)

*EIGEN, L. D. *The Construction of Frames of an Automated Teaching Program.* Collegiate School Automated Teaching Project, New York, N. Y.: November 1959.

*EIGEN, L. D., and KOMOSKI, P. K. Research Summary No. 1 of the Collegiate School Automatic Teaching Project. Report to the Foundation for the Advancement of Education, June 1960. (Mimeo.) U

*EMESON, D. L., and WULFF, J. J. "The Relationship Between 'What Is Learned' and 'How It's Taught.'" Technical Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., December 1957. (ML-TM-57-32) △ L

ENGLISH, H. B. "How Psychology Can Facilitate Military Training—A Concrete Example." *Journal of Applied Psychology* 26: 3-7; February 1942.

ESTES, W. K., and others. *Modern Learning Theory. A Critical Analysis of Five Examples.* New York: Appleton-Century-Crofts, 1954.

EURICH, A. C. "Better Instruction with Fewer Teachers." *Current Issues in Higher Education, 1956.* Washington, D. C.: Association for Higher Education, a department of the National Education Association, 1956. p. 10-16.

EVANS, J. L. "Conversion to Number Bases Other Than Ten." Unpublished programmed text. (Pittsburgh, Pa.: University of Pittsburgh, Department of Psychology, August 1958.) U

*EVANS, J. L. *An Investigation of "Teaching Machine" Variables Using Learning Programs in Symbolic Logic.* Doctor's thesis. University of Pittsburgh, 1960. (Typewritten) [Also in press (mimeo.), to be issued as a report under a grant to the University of Pittsburgh from the Cooperative Research Program of the U. S. Office of Education.] U

**EVANS, J. L.; GLASER, R.; and HOMME, L. E. "A Preliminary Investigation of Variation in the Properties of Verbal Learning Sequences of the 'Teaching Machine' Type." Paper read at the annual convention of the Eastern Psychological Association, Atlantic City, N. J., April 1959. [p. 446-51]

*EVANS, J. L.; HOMME, L. E.; and GLASER, R. *The Ruleg (Rule-Example) System for the Construction of Learning Programs.* A report prepared under the Cooperative Research Program of the U. S. Office of Education at the Department of Psychology, University of Pittsburgh, 1960. (In press, mimeo.)

*EVANS, J. L.; HOMME, L. E.; GLASER, R.; and STELTER, C. *Descriptive Statistics: A Programmed Textbook.* (Vol. I of a Program in Statistics.) Albuquerque, N. M.: Teaching Machines, Inc., 1960.

F

*FATTU, N. A. *A Catalogue of Troubleshooting Tests.* Research Report No. 1, prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, by the Institute of Educational Research, Indiana University, Bloomington, December 1956. △ L

*FATTU, N. A. "Training Devices." *Encyclopedia of Educational Research,* Third edition. New York: The Macmillan Co., 1960.

**FERSTER, C. B., and SAPON, S. M. "An Application of Recent Developments in Psychology to the Teaching of German." *Harvard Educational Review* 28: 58-69; Winter 1958. [p. 173-85]

FERSTER, C. B., and SKINNER, B. F. *Schedules of Reinforcement.* New York: Appleton-Century-Crofts, 1957.

*FINCH, G. (editor) *Symposium on Education and Training Media.* Washington, D. C.: National Academy of Sciences—National Research Council, 1960. (In press).

*Finch, G., and Cameron, F. (editors). *Symposium on Air Force Human Engineering, Personnel, and Training Research.* Washington, D. C.: National Academy of Sciences—National Research Council. Publication 455, 1956. Library of Congress Card Catalog No. 56-60063. (Also identified as ARDC Technical Report 56-8, by the Air Research and Development Command.)

*Finch, G., and Cameron, F. (editors). *Symposium on Air Force Human Engineering, Personnel, and Training Research.* Washington, D. C.: National Academy of Sciences—National Research Council, Publication 516, 1958. Library of Congress Card Catalog No. UG633S96.

Finn, J. D. "A Look at the Future of Audio-Visual Communication." *Audio-Visual Communication Review* 3: 244-56; Fall 1955.

*Finn, J. D. "Automation and Education: I. General Aspects." *Audio-Visual Communication Review* 5: 343-60; Winter 1957. (a)

*Finn, J. D. "Automation and Education: II. Automatizing the Classroom—Background of the Effort." *Audio-Visual Communication Review* 5: 451-67; Spring 1957. (b)

Finn, J. D. "From Slate to Automation." *Audio-Visual Instruction* 4: 84-85, 100-101; March 1959.

**Finn, J. D. "Automation and Education: III. Technology and the Instructional Process." *Audio-Visual Communication Review* 8: 5-26; Winter 1960. [Abridged version of this article appears in this volume under the title "Technology and the Instructional Process." p. 382-94.]

Fisher, R. A. *The Design of Experiments.* Fourth edition. London, England: Oliver & Boyd, 1947.

Flynt, R. C. M., and Hollis, E. V. "Use of Training Aids by Army and Navy." *Higher Education* 2: 1-3; September 1945.

Ford Foundation, Fund for the Advancement of Education. *Teaching by Television.* New York: the Foundation, May 1959.

Freeburne, C. M. "The Influence of Training in Perceptual Span and Perceptual Speed Upon Reading Ability." *Journal of Educational Psychology* 40: 321-52; October 1949.

Freeman, F. N. "An Experiment in the Use of Typewriters in the Elementary School." *Elementary School Journal* 32: 752-59; June 1932.

*Freeman, J. T. "The Effects of Reinforced Practice on Conventional Multiple Choice Tests." *Automated Teaching Bulletin* 1: 19-20; September 1959.

*French, R. S. "Evaluation of a K-System Troubleshooting Trainer." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication 455, 1956. (Library of Congress Card Catalog No. 56-60063.) p. 160-65.

*French, R. S. and Martin, L. B. *A Flight-Line Troubleshooting Trainer for a Complex Electronic System: The "MAC-II Trainer."* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, July 1957. Development Report AFPTRC-TN-57-106. (ASTIA Document No. 134227.) △

*Fry, E. B. "Teaching Machine Dichotomy: Skinner Versus Pressey." Paper read at the annual convention of the American Psychological Association, Cincinnati, September 1959. (In press, *Psychological Reports.*) (a)

*Fry, E. B. "Teaching Machines: An Investigation of Constructed Versus Multiple-choice Methods of Response." *Automated Teaching Bulletin* 1: 11-12; December 1959. (b)

*Fry, E. B. "Teaching Machines: The Coming Automation." *The Phi Delta Kappan* 41: 28-31; October 1959. (c)

**Fry, E. B. "A Study of Teaching Machine Response Modes." (1960) [p. 469-74] (a)

*Fry, E. B. *Teaching Machines: An Investigation of Constructed Versus Multiple-Choice Response Modes.* Doctor's thesis. Los Angeles: University of Southern California, 1960. (b)

*Fry, E. B.; Bryan, G. L.; and Rigney, J. W. "Teaching Machines: An Annotated Bibliography." *Audio-Visual Communication Review* 8: 1-80; Supplement 1, 1960. 80 p.

G

Gagne, R. M. "The Effect of Sequence of Presentation of Similar Items on the Learning of Paired Associates." *Journal of Experimental Psychology* 40: 61-73; February 1950.

Gagné, R. M. "Training Devices and Simulators: Some Research Issues." *American Psychologist* 9: 95-107; March 1954.

Gagné, R. M. "AFPTRC Maintenance Laboratory Research and Development Program." Unpublished Staff Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., December 1957. △ U

*Gagné, R. M. "Teaching Machines and Transfer of Training." Paper read at Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Convention, Cincinnati, September 1959.

*Gagné, R. M., and Bolles, R. C. "A Review of Factors in Learning Efficiency." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 2, p. 13-54. [This paper was earlier given limited distribution as a separate report prepared for the Air Force Office of Scientific Research, November 1958; AFOSR-TN-58-924, ASTIA Document No. AD 162275.]

*Galanter, E. H. (editor) *Automatic Teaching: The State of the Art.* New York: John Wiley & Sons, 1959. (a)

*Galanter, E. H. "Chairman's Summary of Conference." Abstract presented at Air Force Office of Scientific Research and the University of Pennsylvania Conference on the Automatic Teaching of Verbal and Symbolic Skills, Philadelphia, December 1958. (Mimeo.) U

*Galanter, E. H. "The Ideal Teacher." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 1, p. 1-12. (b)

Garver, F. M., and Matthews, R. D. "An Analysis of the Results of Speed Drills with the Metron-o-scope to Increase Reading Rate." *Journal of Educational Psychology* 30: 693-98; December 1939.

*Gates, A. I. "Recitation as a Factor in Memorizing." *Archives of Psychology* 7: 40; 1917.

Gates, A. I. *The Improvement of Reading.* Third edition. New York: Macmillan Co., 1947.

Gaudin, L. S. "The Language Laboratory." *French Reader* 25: 284-91; 1952.

General Dynamics Corporation, Electric Boat Division, (Groton, Conn.) "The All Electronic 'Autoscore.'" Unpublished memorandum. U

Ghiselli, E. E., and Brown, C. W. *Personnel and Industrial Psychology.* New York: McGraw-Hill Book Co., 1948.

Gibson, J. J., editor. *Motion Picture Testing and Research.* Army Air Forces Report, No. 7. Washington, D. C.: Superintendent of Documents, Government Printing Office, 1947.

Gibson, J. J., and others, "History, Organization, and Research Activities, Psychological Test Film Unit, Army Air Force." *Psychological Bulletin* 41: 457-68; July 1944.

Gilbert, A. C. F. "Effect of Immediacy of Knowledge of Correctness of Response Upon Learning." *Journal of Educational Psychology* 47: 415-23; November 1956.

*Gilbert, T. F. "An Early Approximation of Principles of Analyzing and Revising Self-Instructional Programs." Paper read at the Air Force Office of Scientific Research and University of Pennsylvania Conference on the Automated Teaching of Verbal and Symbolic Skills, Philadelphia, December 1958. U (a)

*Gilbert, T. F. "An Early Approximation to Principles of Programming Continuous Discourse, Self-Instructional Materials." A Report to Bell Telephone Laboratories, September 1958. (Unpublished paper, mimeo.) U (b)

*Gilbert, T. F. "Some Issues Involved in Selecting an Automatic Teaching Device." A Report to the Bell Telephone Laboratories, September 1958. (Unpublished paper, mimeo.) U (c)

*Gilbert, T. F. *Some Recent Attempts at the Partial Automation of Teaching.* A Report to the University of Georgia, 1958. (Unpublished paper, mimeo.) U (d)

*Gilbert, T. F. "An Early Approximation to Principles of Programming Continuous Discourse, Self-Instructional Materials." *Psychological Problems and Research Methods in Mathematics Training.* (Edited by R. L. Feierabend and P. H. DuBois.) Washington University, Cooperative Research Project No. 642, USOE, June 1959. (Mimeo.) (a)

**Gilbert, T. F. "On the Relevance of Laboratory Investigation of Learning to Self-Instructional Programming." Symposium paper presented at the Annual Convention of the American Psychological Association, Cincinnati, September 1959. [p. 475-85] (b)

Glaser, R. "Knowledge of Learning and Methods of Training." Unpublished paper. (Pittsburgh, Pa.: University of Pittsburgh, February 28, 1958.) U

**Glaser, R. "Christmas Past, Present, and Future." *Contemporary Psychology* 5: 24-28; 1960. [p. 23-31] (a)

*Glaser, R. "Comments on the Use of Teaching Machine Techniques in the Language Laboratory." *Language Teaching Today.* (Edited by F. J. Oinas.) A Report of the Language Laboratory Conference held at Indiana University, January 22-23, 1960. Publication No. 14 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics. (In press, 1960.) (b)

*Glaser, R. "Principles and Problems in the Preparation of Programmed Learning Sequences." Paper read at the University of Texas Symposium on the Automation of Instruction, University of Texas, May 1960. [Also to be published as a report of a Cooperative Research Program Grant to the University of Pittsburgh under sponsorship of the U. S. Office of Education. In press, 1960.] (c)

**Glaser, R.; Damrin, D. E.; and Gardner, F. M. "The Tab Item: A Technique for the Measurement of Proficiency in Diagnostic Problem Solving Tasks." *Educational and Psychological Measurement* 14: 283-93; Summer 1954. [p. 275-85]

*Glaser, R., and Glanzer, M. *Training and Training Research.* Pittsburgh, Pa.: American Institute for Research, August 1958. △ L

**Glaser, R.; Homme, L. E.; and Evans, J. L. "An Evaluation of Textbooks in Terms of Learning Principles." Paper read at the meeting of the American Education Research Association, Atlantic City, N. J., February 1959. [p. 437-45]

*Glaser, R., and Phillips, J. C. *Guided Missile Personnel Research: Report No. 5, An Analysis of Tests of Proficiency for Guided Missile Personnel: II. The Troubleshooting Board.* Pittsburgh, Pa.: American Institute for Research, 1954.

Glock, M. D. "The Effect Upon Eye Movements and Reading Rate at the College Level of Three Methods of Training." *Journal of Educational Psychology* 40: 93-106; February 1949.

*Goldstein, M., and Chalmers, E. L., Jr. "Photo-Identification Concepts via Teaching Machine." Unpublished paper. (1960) U

Goldstone, R. H. "Experiments with Audio-Visual Aids in Teaching Shakespeare." *College English* 13: 319-22; March 1952.

Grabbe, E.; Ramo, S.; and Wooldridge, D. E. *Handbook of Automation, Computation and Control: Vol. I. Control Fundamentals.* New York: John Wiley & Sons, 1958.

Greenspoon, J., and Foreman, S. "Effect of Delay of Knowledge of Results on Learning a Motor Task." *Journal of Experimental Psychology* 51: 226-28; March 1956.

*Grings, W. W.; Rigney, J. W.; Bond, N. A.; and Summers, S. A. *A Methodological Study of Electronics Troubleshooting Skill: I. Rationale for and Description of the Multiple-Alternative Symbolic Troubleshooting Test.* Technical report prepared for the Personnel and Training Branch, Psychological Sciences Division, Office of Naval Research, at the Department of Psychology, University of Southern California, Los Angeles, August 1953. △ L

Guetzkow, H.; Kelly, E. L.; and McKeachie, W. J. "An Experimental Comparison of Recitation, Discussion, and Tutorial Methods in College Teaching." *Journal of Educational Psychology* 45: 193-207; April 1954.

Guilford, J. P. *Fundamental Statistics in Psychology and Education.* Third Edition. New York: McGraw-Hill Book Co., 1956.

Gustafson, C. B. "A Simple Device for Demonstrating Addition and Subtraction in the Binary Number System." *Mathematics Teacher* 47: 499-500; November 1954.

Guthrie, E. R. *The Psychology of Learning.* New York: Harper & Brothers, 1935.

Guthrie, E. R. "Conditioning: A Theory of Learning in Terms of Stimulus Response Association." *The Psychology of Learning.* Forty-First Yearbook, Part II, of the National Society for the Study of Education. Bloomington, Ill.: the Society, 1942. Chap. 1, p. 17-60.

H

*Hatch, R. S. "An Evaluation of the Effectiveness of a Self-Tutoring Approach Applied to Pilot Training." Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, 1959. WADC Technical Report, TR 59-320. △

Hebb, D. O. *The Organization of Behavior.* New York: John Wiley & Sons, 1949.

Henning, G. "Meeting the Teacher Shortage." Technical Military Planning Operation, General Electric Company, Santa Barbara, Calif., February 1957. U

Hilgard, E. R. *Theories of Learning.* Second Edition. New York: Appleton-Century-Crofts, 1956.

Hilgard, E. R. "Learning Theory and Its Applications." *New Teaching Aids for the American Classroom.* Stanford University, Calif.: The Institute for Communication Research, 1960. p. 19-26. (Library of Congress Card Catalog No. 60-11138)

Hill, W. W. *Tester-Rater, Self-Tutoring Punchboard Device 20-E-2-e.* Port Washington, N.Y.: Special Devices Center, Office of Naval Research, September-October 1955. Special Devices Center Bimonthly Report. U

Hirsch, R. S. *The Effects of Knowledge of Test Results on Learning of Meaningful Material.* (The Pennsylvania State University) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, September 1952. Technical Report SDC 269-7-30. △

Hirsch, R. S. "Mechanical Aids in Language Teaching." *Monograph Series on Languages and Linguistics* 4: 58-63; 1953.

*Hively, W. "Implications for the Classroom of B. F. Skinner's Analysis of Behavior." *Harvard Educational Review* 29: 37-42; Winter 1959.

**Hively, W. "An Exploratory Investigation of an Apparatus for Studying and Teaching Visual Discrimination, Using Pre-school Children." (1960) [p. 247-56]

Hoban, C. F. "The Usable Residue of Educational Film Research." *New Teaching Aids for the American Classroom.* Stanford University, Calif.: The Institute for Communication Research, 1960. p. 95-115. (Library of Congress Card Catalog No. 60-11138)

Hoban, C. F., and Van Ormer, E. B. *Instructional Film Research 1918-1950.* Port Washington, N. Y.: Special Devices Center, Office of Naval Research, 1950. Technical Report SDC 269-7-19. △

*Hoehn, A. J., and Lumsdaine, A. A. *Design and Use of Job Aids for Communicating Technical Information.* Lackland Air Force Base, Tex.: Air Force Personnel and Training Research Center, January 1958. Technical Report AFPTRC-TR-58-7 (ASTIA Document No. 152109). △

Holland, J. G. "Human Vigilance." *Science* 128: 61-67; 1958.

*Holland, J. G. "A Teaching Machine Program in Psychology." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 5, p. 69-82.

**Holland, J. G. "Teaching Machines: An Application of Principles from the Laboratory." (1960) [p. 215-28] [Also published in *Proceedings of the Educational Testing Service Invitational Conference,* October 1959, on "The Impact of Testing on the Educational Process." Princeton, N. J.: Educational Testing Service, 1960. (In press)]

*Holland, J. G., and Skinner, B. F. "A Self-Tutoring Introduction to a Science of Behavior." Unpublished manuscript. (Cambridge, Mass.: Psychological Laboratories, Harvard University, *circa* 1958-59.) U

Hollingworth, H. L. *Educational Psychology.* New York: Appleton-Century-Crofts, 1933.

**Homme, L. E. "The Rationale of Teaching by Skinner's Machines." (1957) [p. 133-36]

**Homme, L. E., and Glaser, R. "Problems in Programming Verbal Learning Sequences." Paper read at Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Convention, Cincinnati, September 1959. [p. 486-96] (a)

*Homme, L. E., and Glaser, R. "Relationships Between the Programmed Textbook and Teaching Machines." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 9, p. 103-108. (b)

Hovland, C. I. "Human Learning and Retention." *Handbook of Experimental Psychology.* (Edited by S. S. Stevens.) New York: John Wiley & Sons, 1951.

Hovland, C. I. "A 'Communication Analysis' of Concept Learning." *Psychological Review* 59: 461-72; 1952.

Hovland, C. I. "A Set of Flower Designs for Experiments in Concept Formation." *American Journal of Psychology* 66: 140-42; 1953.

*Hovland, C. I.; Lumsdaine, A. A.; and Sheffield, F. D. *Experiments on Mass Communication.* Princeton, N. J.: Princeton University Press, 1949.

Hull, C. L. *Principles of Behavior.* New York: Appleton-Century-Crofts, 1943.

I

Imus, H. A.; Rothney, J. W.; and Bear, R. M. *An Evaluation of Visual Factors in Reading.* Hanover, N. H.: Dartmouth College Publications, 1938.

Institute for Applied Experimental Psychology. *Handbook of Human Engineering Data.* Second edition, revised. Medford, Mass.: Tufts College, 1952. SDC 199-1-2a (NAVEXOS P-643) △

Instructional Film Research Program, The Pennsylvania State University. *A Bibliography of Production, Utilization, and Research on Instructional Films.* State College: the Program, 1953. Technical Report SDC 269-7-40. △

*Irion, A. L. "An Investigation of Four Modes of Operation and Three Types of Learning Tasks on the Improved Subject Matter Trainer." Unpublished paper, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., 1957.

*Irion, A. L., and Briggs, L. J. *Learning Task and Mode of Operation Variables in Use of the Subject Matter Trainer.* Lowry Air Force Base, Colo.: Air Force Personnel and Training Research Center, October 1957. Technical Report AFPTRC-TR-57-8 (ASTIA Document No. AD 134252). △

Israel, M. L. "Skinnerian Psychology and Educational Redesign." Paper read at the 1958 Symposium sponsored by the American Psychological Association Committee on the Relations Between Psychology and Education, Washington, D. C., September 1958. (See following reference)

*Israel, M. L. "Educational Redesign: A Proposal." *School Review* 67: 292-304; 1959.

*Israel, M. L. "Variably Blurred Prompting: I. Methodology and Application to the Analysis of Paired-Associate Learning." *Journal of Psychology* 50: 43-52: 1960.

J

Jaspen, N. *Effects on Training of Experimental Film Variables, Study II: Verbalization, "How-It-Works," Nomenclature, Audience Participation, and Succinct Treatment.* (Pennsylvania State College) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, March 1950. Technical Report SDC 269-7-11. △

*Jensen, B. T. "An Independent-Study Laboratory Using Self-Scoring Tests." *Journal of Educational Research* 43: 134-37; October 1949.

*Jones, H. L., and Sawyer, M. O. "A New Evaluation Instrument." *Journal of Educational Research* 42: 381-85; 1949.

Jones, N. F. "A 'Motorized' Reading Project." *English Journal* 40: 313-19; June 1951.

*Jones, R. S. "Integration of Instructional with Self-Scoring Measuring Procedures." *Abstracts of Doctoral Dissertations* 65: 157-65; 1954. (Columbus: The Ohio State University Press.)

K

Kanner, J. H.; Runyon, R. P.; and Desiderato, O. "Television as a Training and Educational Medium." *Audio-Visual Communication Review* 3: 163-72; Summer 1955.

**Keislar, E. R. "The Development of Understanding in Arithmetic by a Teaching Machine." *Journal of Educational Psychology* 50: 247-53; 1959. [p. 425-36] (a)

*Keislar, E. R. "Theoretical Aspects of Automated Teaching." Paper read at the annual convention of the Western Psychological Association, San Diego, Calif., April 1959. U (b)

Keller, F. S., and Schoenfeld, W. N. *Principles of Psychology.* New York: Appleton-Century-Crofts, 1950.

*Kendler, H. H. "Teaching Machines and Psychological Theory." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 15, p. 177-86.

*Kimble, G. A., and Wulff, J. J. *The Effect of "Response Guidance" on the Value of Audience Participation in Training Film Instruction.* Washington, D. C.: Air Force Human Factors Operations Research Laboratories, 1953. (HFORL Report No. 34) △ L

Kinder, J. S. *Audio-Visual Materials and Techniques.* New York: American Book Co., 1950.

Kinder, J. S. "Audio-Visual Research: Where To Find It." *Audio-Visual Communication Review* 1: 234-41; Fall 1953.

Kingsley, H. L. *The Nature and Conditions of Learning.* Second edition. (Revised by R. J. Garry.) Englewood Cliffs, N. J.: Prentice-Hall, 1957.

*Klaus, D. J. (in collaboration with Holmok, E. S.; Lumsdaine, A. A.; Meyer, D. S.; and Quisenberry, K. S.). *High School Physics.* (A two-volume programmed text: Vol. 1, *Introduction and Static Electricity,* 545+p.; Vol. 2, *Reflection and Refraction,* 720+p.) Pittsburgh, Pa.; American Institute for Research, 1960. L

*Klaus, D. J., and Glaser, R. "Studies of Navy Guided Missile Teams." Pittsburgh, Pa.: American Institute for Research, December 1958. △ L

*Klaus, D. J., and Lumsdaine, A. A. *An Experimental Field Test of the Value of Self-Tutoring Materials in High School Physics.* An Interim Report of Progress and Preliminary Findings. Pittsburgh, Pa.: American Institute for Research, April 1960. (Mimeo.) L

*Komoski, K. P. "Origin of Collegiate School Automated Teaching Project." Unpublished report. New York: Collegiate School, 241 W. 77th St., February 1960. U

Kopstein, F. F., and Rockway, M. "Self-Tutoring and Problems in Training." Paper read at the Air Force Office of Scientific Research and the University of Pennsylvania Conference on the Automated Teaching of Verbal and Symbolic Skills, Philadelphia, December 1958. U

Kopstein, F. F., and Roshal, S. M. "Learning Foreign Vocabulary from Pictures Versus Words." Unpublished Staff Research Memorandum, *Training Aids Research Laboratory, AFPTRC,* Air Research and Development Command, Chanute Air Force Base, Ill., November 1954. △ U

*Kopstein, F. F., and Roshal, S. M. "Method of Presenting Word Pairs as a Factor in Foreign Vocabulary Learning." Paper read at the annual convention of the American Psychological Association, San Francisco, Calif., September 1955. U

Kurtz, A. K. *An Application of Learning Theories to the Training of Radio Code Operators.* Technical Report No. 2, prepared for the Office of Naval Research by the Psychological Corporation, New York, 1959. △ L

Kurtz, A. K.; Walter, J. S.; and Brenner, H. R. *The Effects of Inserted Questions and Statements on Film Learning (Rapid Mass Learning).* Port Washington, N. Y.: Special Devices Center, Office of Naval Research, September 1950. Technical Report SDC 269-7-16.

Kurtz, K. H., and Hovland, C. I. "The Effect of Verbalization During Observation of Stimulus Objects upon Accuracy of Recognition and Recall." *Journal of Experimental Psychology* 45: 157-64; March 1953.

L

Lasser, S. "The Impact of Visual Aid Displays Showing a Manipulative Task." *Quarterly Journal of Experimental Psychology* 6: 95-106; November 1954.

Lauer, A. R.; Henry, L. K.; and Fritz, J. F. "Improving the Reading Ability of College Students." (Abstract) *Psychological Bulletin* 35: 670; November 1938.

Learned, W. S., and Wood, B. D. *The Student and His Knowledge.* New York: Carnegie Corp., 1938.

Lewis, D., and Shephard, A. H. "A New Device for Presenting Paired-Associates in Verbal Learning." *American Journal of Psychology* 63: 600-606; October 1950.

Lindquist, E. F. *Design and Analysis of Experiments in Psychology and Education.* Boston: Houghton Mifflin Co., 1953.

Litchfield, M. "Now I Use The Tachistoscope." *Audio-Visual Guide* 19: 25; April 1953.

Litterer, O. F. "An Experimental Study of Visual Apprehension in Reading." *Journal of Applied Psychology* 17: 266-76; June 1933.

**Little, J. K. "Results of Use of Machines for Testing and for Drill upon Learning in Educational Psychology." *Journal of Experimental Education* 3: 45-49; September 1934. [p. 59-65]

Lumsdaine, A. A. "Audio-Visual Research in the U. S. Air Force." *Audio-Visual Communication Review* 1: 76-90; Spring 1953.

*Lumsdaine, A. A. "Possibilities and Implications of Automated Teaching Methods." Paper read at Symposium on Automation of Teaching, at the annual convention of the American Psychological Association, Washington, D. C., August 1958. U

*Lumsdaine, A. A. "Partial and More Complete Automation of Teaching in Group and Individual Learning Situations." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 13, p. 147-66. (a)

*Lumsdaine, A. A. "Response Cueing and 'Size-of-Step' in Automated Learning Programs." Paper read at Symposium on Research Issues in the Study of Human Learning Raised by Developments in Automated Teaching Methods, American Psychological Association Convention, Cincinnati, Ohio, September 1959. U (b)

**Lumsdaine, A. A. "Some Issues Concerning Devices and Programs for Automated Learning." (1959) [p. 517-39] (c)

**Lumsdaine, A. A. "Teaching Machines and Self-Instructional Materials." *Audio-Visual Communication Review* 7: 163-81; Summer 1959 [Reprinted in this volume, with minor changes, under the title "Teaching Machines: An Introductory Overview." (p. 5-22)] (d)

*Lumsdaine, A. A. "Design of Training Aids and Devices." *Human Factors Methods for Systems Design.* (Edited by J. D. Folley, Jr.) Pittsburgh: American Institute for Research, 1960. L (a)

*Lumsdaine, A. A. "The Development of Teaching Machines and Programmed Self-Instruction." *New Teaching Aids for the American Classroom.* Stanford University, Calif.: The Institute for Communication Research, 1960. p. 136-73. (b)

*Lumsdaine, A. A. "Use of Self-Instructional Devices." Human Factors Methods for System Design. (Edited by J. D. Folley.) Pittsburgh, Pa.: American Institute for Research, 1960. p. 291-326. L (c)

Lumsdaine, A. A.; Butler, F. C.; and Meek, C. F. "A Prototype Magazine Loading Sound Film Viewer." Unpublished staff memorandum. Maintenance Laboratory, AFPTRC, 1958. △ U

*Lumsdaine, A. A., and Gladstone, A. I. "Overt Practice and Audio-Visual Embellishments." *Learning from Films.* (Edited by Mark A. May and A. A. Lumsdaine.) New Haven, Conn.: Yale University Press, 1958. p. 58-71.

Lumsdaine, A. A., and Orr, D. B. *Preliminary Evaluative Studies of Science Telecasts.* Pittsburgh, Pa.: American Institute for Research, 1958. △ L

M

McConnell, J. V.; Cutler, R. L.; and McNeil, E. B. "Subliminal Stimulation: An Overview." *American Psychologist* 13: 229-42; May 1958.

McGeoch, J. A. *The Psychology of Human Learning: An Introduction.* New York: Longmans, Green & Co., 1942.

McGeoch, J. A., and Irion, A. L. *The Psychology of Human Learning.* Second edition. New York: Longmans, Green & Co., 1952.

McKown, H. C., and Roberts, A. B. *Audio-Visual Aids to Instruction.* Second edition. New York: McGraw-Hill Book Co., 1949.

McLuhan, M. "What Fundamental Changes Are Foreshadowed in the Prevailing Patterns of Educational Organization and Methods of Instruction by the Revolution in Electronics?" *Current Issues in Higher Education, 1959.* Washington, D. C.: Association for Higher Education, a department of the National Education Association, 1959. p. 176-82.

*McNeil, J. D. "The Great Didactic and Automated Teaching." Unpublished paper. (Los Angeles: University of California, 1960.) U

McNemar, Q. *Psychological Statistics.* New York: John Wiley & Sons, 1949.

*Maccoby, N. and Sheffield, F. D. "Theory and Experimental Research on the Teaching of Complex Sequential Procedures by Alternate Demonstration and Practice." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council, Publication No. 516, 1958. (Library of Congress Card Catalog No. UG 633.S96) p. 99-107.

*Mager, R. F. "Preliminary Studies in Automated Teaching." *Proceedings of the Institute of Radio Engineers,* October 1959. p. 1790. (a)

*Mager, R. F. "Teaching: Today and Tomorrow." *IRE Student Quarterly,* September 1959. p. 31-37. (b)

Manolakes, G. "The Effects of Tachistoscopic Training in an Adult Reading Program." *Journal of Applied Psychology* 36: 410-12; December 1952.

Margolius, G., and Maccoby, N. "Methods of Combining Practice with Filmed Demonstration in Teaching Complex Response Sequences." Paper read at the annual convention of the American Psychological Association, Chicago, Ill., 1956. U

*Mathieu, G. "Automated Language Instruction: A New Deal for Student and Teacher." *Automated Teaching Bulletin* 1: 5-9; December 1959.

*May, M. A., and Lumsdaine, A. A. *Learning from Films.* New Haven, Conn.: Yale University Press, 1958.

Mayer, S. R. *A Method for Studying Aspects of Job Design and Training Requirements of Operators in Computerized Air Defense Systems.* Technical Memorandum, Operational Applications Laboratory, Air Force Cambridge Research Center, Bedford, Mass., January 1959. (OAL-TM-59-1) △ L

*Mayer, S. R., and Westfield, R. L. *A Field Tryout of a Teaching Machine for Training in SAGE Operations.* Technical Memorandum, Operational Applications Laboratory, Air Force Cambridge Research Center, Bedford, Mass., October 1958. (OAL-TM-58-16) △ L

**Mayhew, D. J., and Johnson, A. F. "Teaching Machines." *Journal of Engineering Education* 50: No. 1, 51-52; 1959. [p. 422-24]

**Mellan, I. "Teaching and Educational Inventions." *Journal of Experimental Education* 4: 291-300; March 1936. [p. 265-74]

*Melton, A. W. "The Science of Learning and the Technology of Educational Methods." *Harvard Educational Review* 29: 96-106; Spring 1959. (a)

*Melton, A. W. "Some Comments on 'The Impact of Advancing Technology on Methods in Education' by Dr. Simon Ramo." September 1959. U (b)

Menger, K. "New Approach to Teaching Intermediate Mathematics." *Science* 127: 1320-23; June 6, 1958.

*Meyer, S. R. "A Program in Elementary Arithmetic: Present and Future." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 6, p. 83-84.

**Meyer, S. R. "Report on the Initial Test of a Junior High-School Vocabulary Program." (1960) [p. 229-46]

Michael, D. N. (and Maccoby, N.). *Some Factors Influencing the Effects of Audience Participation on Learning from a Factual Film.* Memorandum Report, Human Resources Research Laboratory, Headquarters Command, Bolling Air Force Base, Washington, D. C., December 1951. HRRL-MR-13A. △ L

*Michael, D. N., and Maccoby, N. "Factors Influencing Verbal Learning from Films Under Conditions of Audience Participation." *Journal of Experimental Psychology* 46: 411-18; December 1953.

*Miller, N. E., and others. "Graphic Communication and the Crisis in Education." *Audio-Visual Communication Review* 5: 1-120; Summer 1957. (special issue)

*Miller, R. B. *Handbook on Training and Training Equipment Design.* Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, 1953. WADC Technical Report 53-136. △

*Miller, R. B. *Psychological Considerations for the Design of Training Equipment.* Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, December 1954. WADC Technical Report No. 54-563. △

Miller, R. B. *A Study of Methods for Determining Conceptual Content of Technical Jobs for Training Purposes.* Pittsburgh, Pa.: American Institute for Research, December 1955. △ L

Miller, R. B. *A Suggested Guide to Functional Characteristics of Training and Training Equipment.* Technical Memorandum, Maintenance Laboratory, AFPTRC, Air Research and Development Command, May 1956. (ML-TM-56-14.) △ L (a)

Miller, R. B. *Task and Part-Task Trainers and Training.* Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, 1956. WADC Technical Report 56-41. △ (b)

Mitchell, M. B. *Education—A New Era Begins.* Reprint of a speech. Chicago: Encyclopaedia Britannica Films, May 1958.

Morley, F. P. "A Teaching-Learning Device for Mechanical Drawing." *Industrial Arts and Vocational Education* 41: 233-34; September 1952.

Morsh, J. E.; Stannard, A. F. B.; and Graham, C. R. "Studies in International Morse Code. V. The Vimy Adaptations of the Code-Voice Method for Initial Instruction in Morse Code Receiving." *Canadian Journal of Psychology* 2: 42-44; March 1948.

*Morton, F. R. "The Language Laboratory as a Teaching Machine." *Language Teaching Today.* (Edited by F. J. Oinas.) A report of the Language Laboratory Conference held at Indiana University, January 22-23, 1960. Publication No. 14 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics. (In press, 1960.)

Mowry, H. W., and Webb, W. B. "Facilitation of Learning by the Classroom Communicator." *Studies in Air Technical Training.* (Edited by H. W. Mowry, W. B. Webb, and E. A. Garvin.) Technical Report, prepared for the Office of Naval Research by the Department of Psychology, Washington University, St. Louis, Mo., December 1955. (Technical Report No. 2, Nonr. 816 (02). △ L

Muller, H. J. *The Uses of the Past.* New York: Oxford University Press (Galaxy), 1957.

Mumford, L. *In the Name of Sanity.* New York: Harcourt, Brace and Co., 1954.

Mumford, L. "An Appraisal of Lewis Mumford's *Technics and Civilization*" (1934); *Daedalus,* Summer 1959 (issued as Vol. 88, No. 3 of the Proceedings of the American Academy of Arts and Sciences). p. 527-36.

Munn, N. L. "Learning in Children." *Manual of Child Psychology.* Second edition. New York: John Wiley & Sons, 1954. p. 374-458.

N

National Education Association, Department of Audio-Visual Instruction. *Television in Instruction: An Appraisal.* Washington, D. C.: the Department, 1958.

Nelson, C. W. *The Multiple Purpose Self-Trainer.* Park Forest, Ill.: Management Research Associates, 1954.

Neville, L. E. "Split-Second Recognition: A New Allied Weapon." *Aviation* 42: 119, 382-85; 1943.

*Newman, S. E. "Student Versus Instructor Design of Study Method." *Journal of Educational Psychology* 48: 328-33; October 1957.

O

O'Connor, V. J. "An Examination of Instructional Films for Characteristics of an Effective Teaching Presentation." *Harvard Educational Review* 20: 271-84; Fall 1950.

Olander, C. "A Model for Visualizing the Formula for the Area of a Circle." *Mathematics Teacher* 48: 245-46; April 1955.

Oliver, J. W. *History of American Technology.* New York: Ronald Press, 1956.

O'Neill, J. C. "Men, Machines and Conversation Courses." *French Review* 26: 439-45; May 1953.

Orbison, W. D. *The Relative Efficiency of Whole and Part Methods of Learning Paired-Associates as a Function of the Length of List."* Doctor's thesis. New Haven, Conn.: Yale University, 1944.

P

*Pask, G. "Automatic Teaching Techniques." *British Communications and Electronics* 4: 210-11; 1957. (a)

Pask, G. "Report on Symposium Papers at Conference on Automata Held in Tedingham, England." *Automation Progress* 4: (No. 2); *circa* 1957. (b)

*Pask, G. "A Teaching Machine for Radar Training." *Automation Progress* 2: 214ff; May 1957. (c)

**Pask, G. "Electronic Keyboard Teaching Machines." *Education and Commerce* 24: 16-26; July 1958. [p. 336-48] (a)

*Pask, G. "Teaching Machines." Unpublished paper. Solartron Electronic Group, Limited, 1958. (Mimeo.) U (b)

*Pask, G. "The Self Organizing Teacher." *Automated Teaching Bulletin* 1 (No. 2): 13-18; December 1959. [See Rheem Manufacturing Company] (a)

*Pask, G. "The Teaching Machine." *The Overseas Engineer,* February 1959. p. 231-32. (b)

Pask, G. "Tomorrow's Control Systems Can Learn from Experience." *Automation Progress* 4: 44-45, 57; February 1959. (c)

**Pask, G. "Adaptive Teaching with Adaptive Machines." (1960) [p. 349-66].

Peins, M. "Mechanical Devices in the Classroom." *Elementary English* 28: 138-40; March 1951.

Pens, E. H. *Explanation of the Functional Block Diagrams for Two Self-Instructional Fire Control System Maintenance Trainers.* Technical Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., December 1957. (ML-TM-57-17) △ L

Perky, C. W. "An Experimental Study of Imagination." *American Journal of Psychology* 21: 422-52; July 1910.

Perry, W. G., Jr., and Whitlock, C. P. "A Clinical Rationale for a Reading Film." *Harvard Educational Review* 24: 6-27; Winter 1954.

*Peterson, J. C. "A New Device for Teaching, Testing, and Research in Learning." *Transactions of the Kansas Academy of Science* 33: 41-47; 1930.

**Peterson, J. C. "The Value of Guidance in Reading for Information." *Transactions of the Kansas Academy of Science* 34: 291-96; 1931. [p. 52-58]

Phillips, J. L., Jr. "Perception in Number Skills: A Study in Tachistoscope Training." *Journal of Educational Psychology* 45: 459-68; December 1954.

*Popp, H., and Porter, D. "Programming Verbal Skills for Primary Grades." Paper read at the annual convention of the Department of Audio-Visual Instruction, NEA, Cincinnati, March 1960. *Audio-Visual Communication Review* (in press).

**Porter, D. "A Critical Review of a Portion of the Literature on Teaching Devices." *Harvard Educational Review* 27: 126-47; Spring 1957. [p. 114-32]

**Porter, D. "A Report on Instructional Devices in Foreign Language Teaching." Prepared for International Training and Research Program of the Ford Foundation, October 1, 1958. (Mimeo.) [p. 186-205] (a)

Porter, D. "A Report on the Use of Audio-Visual Aids in Foreign Language Instruction." A Report to the Fund for the Advancement of Education, April 16, 1958. (Unpublished paper, mimeo.) △ U (b)

**Porter, D. "Teaching Machines." *Harvard Graduate School of Education Association Bulletin* 3: 1-5; March 1958. [p. 206-14] (c)

*Porter, D. "Some Effects of Year Long Teaching Machine Instruction." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 7, p. 85-90.

Pressey, S. L. "The Influence of Color upon Mental and Motor Efficiency." *American Journal of Psychology* 32: 326-56; July 1921.

**Pressey, S. L. "A Simple Apparatus Which Gives Tests and Scores—and Teaches." *School and Society* 23: 373-76; March 20, 1926. [p. 35-41]

**Pressey, S. L. "A Machine for Automatic Teaching of Drill Material." *School and Society* 25: 549-52; May 7, 1927. [p. 42-46]

**Pressey, S. L. "A Third and Fourth Contribution Toward the Coming 'Industrial Revolution' in Education." *School and Society* 36: 668-72; November 19, 1932. [p. 47-51]

*Pressey, S. L. "Further Attempts To Develop a 'Mechanical Teacher.'" Paper read at the annual convention of the American Psychological Association, Philadelphia, September 1946. U [Abstracts in *The American Psychologist* 1 (No. 7); July 1946.]

Pressey, S. L. *Educational Acceleration: Appraisals and Basic Problems.* Bureau of Educational Research Monograph No. 31. Columbus: The Ohio State University, 1949.

**Pressey, S. L. "Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-Instruction." *Journal of Psychology* 29: 417-47; April 1950. [p. 69-88]

*Pressey, S. L. "Some History Regarding Self-Instructional Devices." Paper read at Symposium on the Automation of Teaching, American Psychological Association Convention, Washington, D. C., August 1958.

*Pressey, S. L. "Certain Major Psycho-Educational Issues Appearing in the Conference on Teaching Machines." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 16, p. 187-98.

**Pressey, S. L. "Some Perspectives and Major Problems Regarding 'Teaching Machines.'" (1960) [p. 497-505]

Pressey, S. L., and Hanna, D. C. "The Class As a Psycho-Sociological Unit." *Journal of Psychology* 16: 13-19; July 1943.

Pressey, S. L.; Robinson, F. P.; and Horrocks, J. E. *Educational Psychology.* New York: Harper & Brothers, 1959.

*Price, G. R. "The Teaching Machine." *Think* 25: 10-14; March 1959.

R

**Ramo, S. "A New Technique of Education." *Engineering and Science Monthly* 21: 17-22; October 1957. [p. 367-81]

*Ramo, S. "A New Technique of Education." *Institute of Radio Engineers Transaction on Education* E-1: 37-42; 1958.

*Ramo, S. "The Impact of Advancing Technology on Methods in Education." Paper read at the annual convention of the American Psychological Association, Cincinnati, September 1959.

*Rath, G. J.; Andersen, N. S.; and Brainerd, R. C. "The IBM Research Center Teaching Machine Project." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 11, p. 117-30.

Raymond, R. C. *"The Need for Labor-Saving Machinery in Education.* Reprint of a speech. Technical Military Planning Operation, General Electric Company, Santa Barbara, Calif., February 1957.

Reid, L. S. "The Development of Noncontinuity Behavior Through Continuity Learning." *Journal of Experimental Psychology* 46: 107-12; 1953.

*Reid, L. S. "A Review of *Automatic Teaching: The State of the Art,* E. H. Galanter, Editor." *Science* 131: 29-30; 1960.

Reiss, F. A. "German Syntax on the Screen." *Modern Language Journal* 35: 135-38; February 1951.

*Renner, Inc. *Inexpensive Missile Training—A New Approach.* Paper read at the U. S. Navy Guided Missile Training Symposium, Naval Ordnance Laboratory, Whiteoak, Md., March 1958.

Renshaw, S. "The Visual Perception and Reproduction of Forms by Tachistoscopic Methods." *Journal of Psychology* 20: 217-35; 1945.

*Rheem Califone Corporation. *Automated Teaching Bulletin* 1: 1-28; September 1959.

*Rheem Califone Corporation. *Automated Teaching Bulletin* 2; December 1959.

*Rigney, J. W.; Bryan, G. L.; and Svenson, D. W. "The Multiple Choice Item as a Training Device." Paper read at the annual convention of the American Psychological Association, Chicago, September 1956. U

Ritchie, M. L., and Michael, A. L. "Transfer Between Instrument and Contact Flight Training." *Journal of Applied Psychology* 39: 145-49; June 1955.

Robinson, F. P. "The Tachistoscope as a Measure of Reading Perception." *American Journal of Psychology* 46: 132-35; January 1934.

Robinson, H. M. "Tachistoscope in Reading." *Elementary School Journal* 54: 490-91; May 1954.

Rogers, C. R., and Skinner, B. F. "Some Issues Concerning the Control of Human Behavior: A Symposium." *Science* 124: 1057-1066; November 30, 1956.

Root, R. T. *An Annotated Bibliography of Research on Training Aids and Training Devices.* Staff Memorandum, Training Methods Division, HumRRO, George Washington University, Washington, D. C., August 1957. △ L

Roshal, S. M. *Effects of Learner Representation in Film-Mediated Perceptual-Motor Learning.* (The Pennsylvania State University) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, 1949. Technical Report SDC 269-7-5. △

Roshal, S. M. "The Instructional Film." *USAF-NRC Educational and Training Media: A Symposium.* (Edited by G. Finch.) Washington, D. C.: National Academy of Science—National Research Council, 1960. (In press)

Rothkopf, E. Z. "A Deduction from an Excitation-Inhibition Account of Retroactive Inhibition." *Journal of Experimental Psychology* 53: 207-13; March 1957. (a)

**Rothkopf, E. Z. "Heuristic Discussions of Psychological Bases for the Conduct of Training by Automatic Devices: I. A Functional Diagram and an Approach to Research." Unpublished Staff Memorandum, Maintenance Laboratory, AFPTRC, Lowry Air Force Base, Colo., April 1957. [p. 305-10] (b)

Rothkopf, E. Z. "Signal Similarity and Optimal Signal Arrangement of Lesson and Drill Orders in Alphabetic Morse Code Training." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council Publication 516, 1958. (Library of Congress Card Catalog No. UG633-S96) p. 188-93. (a)

**Rothkopf, E. Z. "Some Research Problems in the Design of Materials and Devices for Automated Teaching." Paper read at Symposium on Automated Teaching at the annual convention of the American Psychological Association, Washington, D. C., August 1958. [p. 318-28] (b)

*Rothkopf, E. Z. "The State of an Art?" (A Review of *Automatic Teaching: The State of the Art*, edited by E. H. Galanter.) *Contemporary Psychology* 5: 104-105; 1960.

S

Saltzman, I. J. "Delay of Reward and Human Verbal Learning." *Journal of Experimental Psychology* 41: 437-39; June 1951.

*Saltzman, I. J. "Comments on Teaching Machines." *Language Teaching Today.* (Edited by F. J. Oinas.) A Report of the Language Laboratory Conference held at Indiana University, January 22-23, 1960. Publication No. 14 of the Indiana University Research Center in Anthropology, Folklore, and Linguistics. (In press, 1960.)

Sax, G. *Concept Acquisition as a Function of Differing Schedules and Delays of Reinforcement.* Doctor's thesis. Los Angeles: University of Southern California, 1958. (Typewritten)

Schaughency, M. D. "Teaching Arithmetic with Calculators." *Arithmetic Teacher* 2: 21-22; February 1955.

*Schutz, R. E. "Progress and Preparation of Automated Instructional Material and Criterion Measures." *Automated Teaching Bulletin* 1:31-33; December 1959 [see Rheem Califone Corporation]. [Also presented at Symposium on Automated Teaching, Western Psychological Association Convention, San Diego, Calif., April 1959.]

*Severin, D. G. "Appraisal of Special Tests and Procedures Used with Self-Scoring Instructional Testing Devices. *Abstracts of Doctoral Dissertations* 66: 323-30; 1955. (Columbus, The Ohio State University Press.)

Shephard, A. H., and Lewis, D. "A New Device for Presenting Paired-Associates in Verbal Learning." *American Journal of Psychology* 63: 600-606; October 1950.

Sidman, M.; Keller, F. S.; Kennedy, E. J.; and Wilson, M. P. "Teaching Morse Code Reception with Signals Weighted in Frequency According to Their Difficulty." *Journal of Applied Psychology* 39: 1-4; February 1955.

*Silberman, H. F. "Introductory Description of Teaching Machines (Physical Characteristics)." Paper read at Symposium on Automated Teaching, Western Psychological Association Convention, San Diego, Calif., April 1959. U

*Silberman, H. F. "The Automation of Teaching." *Behavioral Science,* 1960 (in press).

*Silberman, H. F., and Coulson, J. "A Draft Summary of Findings in an Exploratory Teaching Machine Study." *Automated Teaching Bulletin* 1: (No. 2) 35-37; December 1959.

Skelley, J. F. "Use of the Vu-graph in Remedial Reading." *Audio-Visual Guide* 18: 25; February 1952.

Skinner, B. F. *Science and Human Behavior.* New York: Macmillan Co., 1953. (a)

Skinner, B. F. "Some Contributions of an Experimental Analysis of Behavior to Psychology as a Whole." *American Psychologist* 8: 69-78; February 1953. (b)

**Skinner, B. F. "The Science of Learning and the Art of Teaching." *Harvard Educational Review* 24: 86-97; Spring 1954. [p. 99-113]

Skinner, B. F. "The Experimental Analysis of Behavior." *American Scientist* 45: 343-71; September 1957. (a)

*Skinner, B. F. "Report #1 on the Development of Methods of Preparing Materials for Teaching Machines." Unpublished Report to Human Resources Research Office, January 1957. △ U (b)

*Skinner, B. F. "Report #2 on the Development of Methods of Preparing Materials for Teaching Machines." Unpublished Report to Human Resources Research Office, April 1957. △ U (c)

*Skinner, B. F. "Final Report to Human Resources Research Office." Unpublished report to Human Resources Research Office, October 1957. △ U (d)

Skinner, B. F. *Verbal Behavior.* New York: Appleton-Century-Crofts, 1957. (e)

*Skinner, B. F. "Programming Verbal Knowledge for Automated Instruction in Schools and Colleges." Paper read at Symposium on the Automation of Teaching, American Psychological Association Convention, Washington, D. C., August 1958. U (a)

Skinner, B. F. "Reinforcement Today." *American Psychologist* 13: 94-99; March 1958. (b)

**Skinner, B. F. Teaching Machines." *Science* 128: 969-77; October 24, 1958. [p. 137-58] (c)

*Skinner, B. F. "Comments on 'The Impact of Advancing Technology on Methods in Education' by Dr. Simon Ramo." September 1959. U (a)

*Skinner, B. F. "The Programming of Verbal Knowledge." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 4, p. 63-68. (b)

*Skinner, B. F. "Special Problems in Programming Language Instruction for Teaching Machines." *Language Teaching Today.* (Edited by F. J. Oinas.) A Report of the Language Laboratory Conference held in Indiana University, January 22-23, 1960. Publication No. 14 of the Indiana University Research Center in Anthropology, Folklore, Linguistics. (In press, 1960.)

**Skinner, B. F., and Holland, J. G. "The Use of Teaching Machines in College Instruction." Final Report to The Fund for the Advancement of Education, August 15, 1958. [p. 159-72].

*Smith, D. E. P. "Speculations: Characteristics of Successful Programs and Programmers." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 8, p. 91-102.

Smith, D. E. P.; Wood, R. L.; and Carrigan, P. M. *The SA-S Senior and Junior Scales.* Chicago: Psychotechnics, Inc. (In press.)

Smith, H. P., and Tate, T. R. "Improvements in Reading Rate and Comprehension of Subjects Training with the Tachistoscope." *Journal of Educational Psychology* 44: 176-84; March 1953.

Smith, K. R. and Van Ormer, E. B.. *Learning Theories and Instructional Film Research.* Port Washington, N. Y.: Special Devices Center, Office of Naval Research, June 1949. Technical Report SDC 269-7-6. △

Snyder, A. "The Flashreader in the Reading Laboratory." *English Journal* 41: 269; May 1952.

Speer, G. S. "Using Mechanical Devices Can Increase Speed of Reading." *The Nation's Schools* 48: 45-48; October 1951.

Spence, K. W. "Theoretical Interpretations of Learning." *Handbook of Experimental Psychology.* (Edited by S. S. Stevens.) New York: John Wiley & Sons, 1951.

*Spence, K. W. "The Relation of Learning Theory to the Technology of Education." *Harvard Educational Review* 29: 84-95; Spring 1959.

Standlee, L. S.; Popham, W. J.; and Fattu, N. A. *A Review of Troubleshooting Research.* Research Report No. 3, prepared for the Personnel and Training Branch, Psychological Sciences Division. Office of Naval Research by the Institute of Educational Research, Indiana University, December 1956. △ L

Staton, T. F., and Maize, R. C. "A Voluntary Reading-Improvement Program for Air Force Officers." *School and Society* 76: 42-44; July 19, 1952.

**Stephens, A. L. "Certain Special Factors Involved in the Law of Effect." *Abstracts of Doctoral Dissertations.* (Columbus: The Ohio State University Press.) No. 64, 1953. [p. 89-93]

Stevens, J. M. *Educational Psychology.* Revised edition. New York: Henry Holt and Co., 1956.

Stevens, S. S., editor. *Handbook of Experimental Psychology.* New York: John Wiley & Sons, 1951.

Stoddard, A. J. *Schools for Tomorrow.* New York: The Fund for the Advancement of Education, 1957.

*Stolurow, L. M. "Teaching Machines and Special Education." An unpublished report of work sponsored by the U. S. Office of Education Cooperative Research Project performed at the University of Illinois. 1960. U

Stone, G. R. *The Training Function of Examinations: Retest Performance as a Function of the Amount and Kind of Critique Information.* Lackland Air Force Base, Texas: Air Force Personnel and Training Research Center, June 1955. Research Report AFPTRC-TN-55-8. △

Sueltz, B. A. "Counting Devices and Their Uses." *Arithmetic Teacher* 1: 25-30; February 1954.

Swanson, R. A.; Lumsdaine, A. A.; and Aukes, L. E. "Two Studies in Evaluation of Maintenance Training Devices." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences—National Research Council. Publication 455, 1956. (Library of Congress Card Catalog No. 56-60063) p. 267-75.

Symonds, P., and Chase, D. H. "Practice vs. Motivation." *Journal of Educational Psychology* 20: 19-35; January 1929.

T

Taris, C. M. "Automatic 'Answer Only' Set." *Bell Laboratories Record* 34: 165-68; May 1956.

Taves, M. J., and Gross, N. "A Critique of Rural Sociology, 1950." *Rural Sociology* 17: 109-18; 1952.

Taylor, E. A. *Controlled Reading.* Chicago: University of Chicago Press, 1937.

Thompson, W. C. "A Book Versus Machine Experiment in Adult Reading Improvement." *College English* 15: 471-73; May 1954.

Thorndike, E. L. *Fundamentals of Learning.* New York: Bureau of Publications, Teachers College, Columbia University, 1932.

Tinker, M. A., and Russell, W. A. *Introduction to Methods in Experimental Psychology.* New York: Appleton-Century-Crofts, 1958.

Tolman, E. C. *Purposive Behavior in Animals and Men.* Berkeley, Calif.: University of California Press, 1951. (Originally published in 1932.)

Torreyson, H. C. "Equation Balances." *School Science and Mathematics* 55: 104-108; February 1955.

Traxler, A. E. "Value of Controlled Reading: Summary of Opinion and Research." *Journal of Experimental Education* 11: 280-92; June 1943.

Trump, J. *Images of the Future.* Washington, D. C.: National Association of Secondary-School Principals, a department of the National Education Association, 1959.

Twyford, L. C. "Profile Techniques for Program Analysis." *Audio-Visual Communication Review* 2: 243-61; Fall 1954.

*Twyford, L. C.; O'Hara, J. G.; and Goldman, A. *Green Light Rater.* Port Washington, N. Y.: U. S. Naval Training Device Center, December 23, 1958. 5 p. Technical Report NTDC 20-OS-3-1. △

U

*Underwood, B. J. "Verbal Learning in the Educative Process." *Harvard Educational Review* 29: 107-17; Spring 1959.

U. S. Air Force, Air Force Personnel and Training Research Center. *Index to Air Force Personnel and Training Research Center Technical Documentary Reports* (1957 and 1958). Lackland Air Force Base, Tex.: the Center, March 1958. △

U. S. Department of Defense, Research and Development Board, Committee on Human Resources, Panel on Training and Training Devices. *Symposium on Psychology of Learning Basic to Military Training Problems.* May 7-8, 1953.

U. S. Department of Health, Education, and Welfare. *Annual Report of the U. S. Department of Health, Education, and Welfare: 1953.* Washington, D. C.: Superintendent of Documents, Government Printing Office, 1953.

U. S. Navy, Special Devices Center. *Instructional Film Research Reports.* (The Pennsylvania State University) Port Washington, N. Y.: the Center, 1953. Technical Report SOC 269-7-36; NAVEXOS-P-1220. Volume I: (this report includes SDC 269-7-1 to 35 except 269-7-19) △

U. S. Navy, Naval Training Device Center. *Instructional Film Research Reports.* Vol. II. (The Pennsylvania State University) Port Washington, N. Y.: the Center, June 1956. NTDC Report 269-7-61 (NAVEXOS P-1543). △

U. S. Navy, Naval Training Device Center. *Training Device Guide* [formerly *Special Devices Guide*]. Revised edition. Port Washington, N. Y.: the Center, November 1957. NAVEXOS. △

U. S. NAVY, NAVAL TRAINING DEVICE CENTER. *Bibliography of Unclassified Human Engineering Reports.* Port Washington, N. Y.: the Center, January 1958. △

U. S. NAVY, SPECIAL DEVICES CENTER, OFFICE OF NAVAL RESEARCH. *Instructor's Guide for Tester-Rater, Self-Tutoring Punchboard.* Port Washington, N. Y.: the Center, July 1954. △

UNZICKER, C. E. *An Experimental Study of the Effect of the Use of the Typewriter on Beginning Reading.* Teachers College Contributions to Education, No. 610. New York: Teachers College, Columbia University, 1934.

V

VERPLANCK, W. S. "A Glossary of Some Terms Used in the Objective Science of Behavior." *Supplement to the Psychological Review* 64: November 1957.

VINCENT, W. S.; ASH, P.; and GREENHILL, L. P. *Relationship of Length and Fact Frequency to Effectiveness of Instructional Motion Pictures.* (The Pennsylvania State University) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, 1949. Technical Report SDC 269-7-7. △

VITELES, M. S. *Psychological Principles in the Design and Operation of Synthetic Trainers.* Memorandum 19, Project N-105. Philadelphia: University of Pennsylvania, 1945. L

VOSS, H. A., and RHOADS, C. S. *Application of Electro-Physiological Techniques to Human Performance: The Reading Assessor—The Alertness Indicator.* (Institute for Applied Experimental Psychology, Tufts College) Port Washington, N. Y.: Special Devices Center, Office of Naval Research, 1950. Technical Report SDC-58-2-11. △

VOTAW, D. F., and DANFORTH, L. "The Effect of Method of Response upon the Validity of Multiple-Choice Tests." *Journal of Educational Psychology* 30: 624-27; November 1939.

W

WAGNER, R. W. "The Teaching Film of Tomorrow." *Audio-Visual Communication Review* 2: 216-19; Summer 1954.

WARREN, A. B., and BROWN, R. H. "Conditioned Operant Response Phenomena in Children." *Journal of General Psychology* 28: 181-207; April 1943.

WARREN, N.; ATKINS, D. W.; FORD, J. S.; and WOLBERS, H. L. "Development of a Training Program for Teaching Basic Principles of Trouble Shooting." Technical Memorandum, Armament Systems Personnel Research Laboratory, AFPTRC, Lowry Air Force Base, Colo., October 1955. (ASPRL-TM-55-19) △ L

WASHBURN, J. N. "The Use of Questions in Social Science Material." *Journal of Educational Psychology* 20: 321-59; May 1929.

WEDEEN, S. U. "Mechanical Versus Non-Mechanical Reading Techniques for College Freshman." *School and Society* 79: 121-23; April 17, 1954.

**WEIMER, P. K. "A Proposed 'Automatic' Teaching Device." *Institute of Radio Engineers Transactions on Education* E-1: 51-53; June 1, 1958. [p. 311-17]

WENDT, P. R. "Magnetic Sound Film Has Many Classroom Uses." *The Nations Schools* 51: 88-90; June 1953.

WESTOVER, F. L. *Controlled Eye Movements Versus Practice Exercises in Reading.* Teachers College Contributions to Education, No. 917. New York: Teachers College, Columbia University, 1946.

WIENER, N. *Cybernetics.* New York: John Wiley & Sons, 1948.

WILLEY, C. F. "Classroom Scoring of Tests." *Psychological Reports* 4: 611-17; December 1958.

WITCHER, C. M. "The Teaching of Mathematics to the Blind." *Mathematics Teacher* 48: 314-21; May 1955.

WITTICH, W. A., and SCHULLER, C. F. *Audio-Visual Materials, Their Nature and Use.* New York: Harper & Brothers, 1953.

WITTY, P. A. "Problems in the Improvement and Measurement of Growth in Reading." *School and Society* 78: 69-73; September 5, 1953.

WITTY, P. A. "Evaluation of Methods and Devices to Improve Reading Rate and Comprehension." *Elementary English* 31: 260-67; May 1954.

WITTY, P. A., and FITZWATER, J. P. "An Experiment with Films, Film-Readers, and the Magnetic Sound Track Projector." *Elementary English* 30: 232-41; April 1953.

WITTY, P. A., and KOPEL, D. *Reading and the Educative Process.* Boston: Ginn and Co., 1939.

WITTY, P. A.; STOLARZ, T.; and COOPER, W. "Some Results of a Remedial Reading Program for College Students." *School and Society* 76: 376-80; December 13, 1952.

WITZEMAN, B. E. "An Experimental Study Using the Ophthalm-o-graph and Metron-o-scope in the Diagnosis and Treatment of Reading Defects." *Journal of Psychology* 11: 307-34; April 1941.

WOLFLE, D. *The Use and Design of Synthetic Trainers for Military Training.* Washington, D. C.: Applied Psychology Panel, OSRD, 1946. OSRD Report 5246. △ L

WOLFLE, D. "Training." *Handbook of Experimental Psychology.* (Edited by S. S. Stevens.) New York: John Wiley & Sons, 1951.

WOODWORTH, R. S. "A Contribution to the Question of 'Quick Learning, Quick Forgetting'." *Psychological Bulletin* 11: 58-59; 1914.

WOODWORTH, R. S. *Experimental Psychology.* New York: Henry Holt and Co., 1938.

WOODWORTH, R. S., and SCHLOSBERG, H. *Experimental Psychology.* Revised edition. New York: Henry Holt and Co., 1954.

WOOSTER, G. F. "An Experimental Study of the Reading-Rate Controller." *Journal of Educational Psychology* 45: 421-26; November 1954.

WULFF, J. J. "Specifications for a Paired-Associate Trainer." Unpublished Staff Memorandum, *Maintenance Laboratory, AFPTRC,* Lowry Air Force Base, Colo., December 6, 1957. △ U

WYCKOFF, L. B. "The Role of Observing Responses in Discrimination Learning." *Psychological Review* 59: 431-42; 1952.

Y

YALE MOTION PICTURE RESEARCH PROJECT. "Do Motivation and Participation Questions Increase Learning? *Educational Screen* 26: 256ff; May 1947.

Z

*ZEAMAN, D. "Comments on the Automation of Teaching." Presented at Symposium on the Automation of Teaching, American Psychological Association Convention, Washington, D. C., August 1958.

*ZEAMAN, D. "Skinner's Theory of Teaching Machines." *Automatic Teaching: The State of the Art.* (Edited by E. H. Galanter.) New York: John Wiley & Sons, 1959. Chap. 14, p. 167-76.

INDEX

A

Adams, Henry 385
Adkins, D. C. 581
Allen, W. H. 17, 581, 666
Amsel, A. 399, 506, 582
Anderson, I. H. 115
Anderson, Nancy S. 20, 288, 532, 628, 676
Angell, D. 582, 604
Angell, George W. 34, 66, 120, 125, 127, 583, 644, 672
Arnoult, M. D. 583
Ash, P. 581
Ashby, Ross 390
Auble, D. 123, 124
Azrin, Nathan H. 154, 219

B

Bailey, C. 341
Ball, J. 602
Barbe, W. B. 115
Barlow, John A. 396, 416, 583, 584, 628
Bartlett, F. C. 349
Beck, J. 549, 553, 584, 585, 628, 656
Berg, I. A. 124
Berger, E. J. 115
Bertone, C. M. 452
Besnard, G. G. 259, 261, 274, 299, 300, 302, 585, 591, 621, 622
Bijou, S. W. 125
Binet, Alfred 500
Blough, Donald S. 102, 103
Blyth, J. W. 396, 401, 585
Bolles, R. C. 627, 628
Bond, N. A., Jr. 594, 638
Bortolazzo, J. L. 118
Brainerd, R. C. 288, 628, 676
Brandwein, P. F. 499
Brethower, D. M. 221
Brewster, J. R. 115
Briggs, Leslie J. 6, 10, 20, 34, 258, 259, 261, 262, 299, 300, 302, 305, 321, 329, 335, 525, 544, 545, 561, 585, 586, 587, 588, 589, 590, 591, 592, 604, 621, 622, 628, 642, 650, 675, 688
Brown, G. I. 592
Brown, J. S. 259, 597, 621
Brown, R. H. 123
Bryan, G. L. 2, 259, 575, 594, 595, 596, 597, 608, 609, 610, 621, 625, 677
Buitenkand, N. 597
Byrnes, F. C. 602

C

Cameron, F. 258, 591, 602, 614, 622, 624, 655
Cantor, J. H. 259, 597, 621
Carlton, B. J. 581
Carney, C. D. 115
Carpenter, C. R. 598
Carpenter, F. 598
Carr, W. J. 400, 540, 600
Carroll, J. B. 180, 182
Carter, H. L. 115
Chalmers, E. L., Jr. 638
Chard, H. 265, 269
Chase, D. H. 57
Comenius, John Amos 656
Composto, Frank R. 274
Cook, D. A. 600
Cook, J. O. 259, 547, 582, 601, 602, 603, 604, 622, 666
Cooney, S. 452
Cooper, Debra 154
Cornell, F. G. 275, 605
Corrigan, R. E. 605, 677
Coulson, J. E. 397, 426, 452, 534, 543, 605, 606, 607, 677, 681
Counts, G. S. 114
Cross, K. P. 607
Crowder, Norman A. 20, 21, 258, 259, 260, 261, 286, 305, 356, 389, 426, 505, 526, 528, 532, 533, 542, 549, 558, 587, 589, 590, 605, 606, 608, 609, 610, 628, 629, 634, 638, 650, 653, 654
Curtis, F. D. 68, 128

D

Dallenbach, K. M. 115
Damrin, Dora E. 260, 275, 605, 610, 637, 648
Danforth, L. 126
Darby, C. L. 2, 608, 610
David, E. E., Jr. 193
Davis, H. M. 274
Day, J. H. 610
Dear, R. E. 452
Dearborn, W. F. 115, 116
Dewey, John 158
Diederich, P. B. 216, 610
Dowell, E. C. 611
DuBois, P. H. 636
DuVall, W. E. 262, 588, 591

E

Ebbinghaus, H. 475, 483
Eckstand, Gordon 666
Edison, Thomas A. 386
Edwards, A. L. 450
Edwards, Ward D. 96, 118, 507, 611, 612, 613, 614, 622
Eigen, Lewis D. 615
Ellis, W. R., Jr. 452
Ellson, D. G. 278
Emeson, D. L. 616, 617
English, H. B. 122
Esgate, K. 452
Estes, W. K. 321, 654
Evans, James L. 12, 17, 389, 397, 437, 446, 447, 494, 534, 543, 546, 549, 556, 557, 558, 559, 560, 561, 618, 619, 620, 637, 641, 647, 649

F

Fattu, N. A. 259, 260, 594, 620, 621
Feierabend, R. L. 636
Ferster, Charles B. 96, 100, 102, 136, 154, 156, 170, 173, 419, 425, 470, 506, 561, 592, 621, 630
Finch, G. 258, 589, 591, 602, 614, 621, 622, 624, 655
Finn, James D. 264, 382, 622, 623
Folley, J. D., Jr. 651, 655
Foster, R. R. 55
Freeman, J. T. 623, 624, 677
French, R. S. 259, 622, 624
Friberg, Mrs. Carl 247
Fry, Edward B. 2, 397, 398, 469, 505, 543, 575, 591, 595, 596, 597, 598, 605, 607, 608, 609, 610, 611, 615, 621, 624, 625, 626, 628, 639, 640, 643, 645, 649, 657, 658, 665, 667, 670, 672, 674, 675, 676, 677, 678, 680, 686, 687
Fuller, C. 452

G

Gagné, R. M. 20, 261, 322, 547, 626, 627, 628
Gaier, E. L. 607
Galanter, Eugene H. 2, 155, 497, 584, 588, 609, 627, 628, 640, 641, 646, 650, 664, 673, 676, 677, 678, 685, 686, 689
Gardner, Floyd M. 260, 275, 605, 637, 648
Gates, A. I. 131, 629
Gaudin, L. S. 131
Gibson, J. J. 118, 581
Gilbert, Thomas F. 17, 155, 156, 195, 198, 398, 475, 486, 490, 540, 541, 543, 544, 545, 546, 549, 552, 556, 557, 558, 560, 629, 630, 631, 632, 633, 634, 635, 636, 656, 685
Gladstone, A. I. 655, 657, 665
Glanzer, M. 259, 637
Glaser, Robert 3, 12, 13, 15, 17, 23, 229, 258, 259, 260, 263, 275, 369, 397, 399, 426, 437, 446, 447, 453, 486, 533, 534, 542, 543, 546, 549, 556, 557, 558, 559, 560, 561, 605, 618, 619, 620, 621, 628, 636, 637, 638, 641, 647, 648, 649, 659
Godcharles, Charles 401
Goldman, A. 688
Goldstein, M. 638
Graham, C. R. 320
Grings, W. W. 594, 621, 638
Gross, N. 118
Guetzkow, H. 127
Guilford, J. P. 473
Gustafson, C. B. 115
Guthrie, E. R. 67, 94, 115, 122, 437, 517, 654

H

Hatch, R. S. 262, 638
Henderson, T. E. 452
Herold, E. W. 311
Hilgard, E. R. 91, 497
Hively, Wells 97, 98, 217, 247, 639
Hoban, C. F. 118, 194, 259, 657, 666
Hodgkinson, H. L. 592
Hoehn, A. J. 17, 639
Holland, James G. 16, 96, 97, 137, 154, 155, 159, 215, 221, 222, 226, 234, 247, 397, 418, 419, 458, 628, 639, 640, 656, 675, 686
Holmok, E. S. 647
Holt, W. R. 274
Homme, Lloyd E. 12, 13, 15, 17, 96, 133, 154, 229, 389, 397, 399, 426, 437, 446, 447, 453, 486, 533, 534, 542, 543, 546, 549, 556, 557, 558, 559, 560, 561, 618, 619, 620, 628, 637, 641, 647, 649
Hovland, C. I. 223, 259, 561, 581, 629, 641, 651, 655, 665
Hull, C. L. 67, 122, 123, 129
Huxley, Aldous 664

I

Irion, A. L. 261, 299, 302, 544, 545, 561, 604, 642
Israel, M. L. 98, 539, 643, 644

J

Jacobson, John 404
Jaspen, N. 581
Jensen, B. T. 34, 120, 128, 129, 452, 501, 644, 675
Jewett, J. E. 155
Johnson, A. F. 396, 422, 658
Jones, D. L. 591
Jones, H. L. 34, 644
Jones, R. S. 34, 502, 504, 644, 645, 675
Judd, C. H. 427

K

Keislar, Evan R. 396, 425, 505, 548, 556, 557, 561, 645, 646
Keller, F. S. 221, 320
Kelley, E. L. 127
Kendler, H .H. 557, 628, 646
Kendler, T. S. 547, 582, 601, 602, 604, 622
Kennedy, E. J. 320
Kimble, G. A. 17, 259, 604, 646, 647, 651
Kinder, J. S. 118
Klaus, D. J. 259, 569, 647, 648
Komensky (Comenius), Jan Amos 656
Komoski, K. P. 615, 649
Kopstein, F. F. 547, 649
Kurtz, K. H. 581

L

LaPorte, H. R. Jr. 594
Learned, W. S. 500
Lewis, D. 307
Lindquist, E. F. 55
Lindsley, Ogden R. 103
Little, James Kenneth 33, 59, 66, 125, 127, 129, 130, 501, 504, 650, 675
Lumsdaine, A. A. 3, 5, 17, 258, 259, 261, 294, 300, 400, 517, 523, 543, 561, 569, 581, 582, 587, 604, 628, 634, 639, 641, 643, 647, 648, 650, 651, 653, 655, 657, 665, 689

Mc

McGeoch, J. A. 125, 322
McKeachie, W. J. 127

McKown, H. C. 118
McLuhan, M. 390, 393
McNeil, J. D. 656

M

Maccoby, N. 17, 259, 533, 548, 581, 622, 639, 651, 655, 665
Mager, R. F. 657
Mann, Horace 498
Margolius, G. 533
Martin, L. B. 624
Mathes, J. A. 274
Mathieu, G. 657, 677
May, Mark A. 259, 655, 657
Mayer, S. R. 262, 658
Mayhew, D. J. 396, 422, 658
Mech, E. V. 123, 124
Mellan, Ibert 257, 265, 658
Melton, A. W. 263, 658, 659, 660, 676
Menger, K. 157
Meyer, D. S. 647
Meyer, Susan R. 97, 134, 145, 154, 229, 549, 555, 556, 558, 628, 649, 664, 665
Michael, D. N. 548, 581, 651, 665
Miesegaes, W. 274
Miller, N. E. 657, 665
Miller, R. B. 259, 666, 667
Mitchell, Maurice 389
Montessori, Maria 265
Morley, F. P. 115
Morsh, J. E. 320
Morton, F. R. 667
Muller, Herbert J. 388, 394
Mumford, L. 388
Munn, N. L. 123
Mursch, G. A. 300, 585

N

Newman, S. E. 617, 667

O

O'Connell, S. L. 452
O'Hara, J. G. 688
Ohm, Georg Simon 356
Oinas, F. J. 636, 667, 676, 686
Olander, C. 115
Oliver, J. W. 385
Orbison, W. D. 322, 323
Osgood, C. E. 308

P

Parkinson, C. D. 452
Pask, Gordon 19, 262, 263, 319, 320, 336, 349, 350, 356, 366, 390, 391, 530, 536, 668, 669, 670, 672, 677
Patton, B. 452
Peins, M. 132
Perry, W. G., Jr. 115, 130
Peterson, H. J. 9, 33, 51, 52, 68, 88, 120, 127, 129
Peterson, J. C. 33, 52, 66, 88, 120, 127, 129, 672, 675
Phillips, J. C. 260, 621, 637
Plashinski, D. 591
Poehler, Paul 247
Popp, H. 672
Poritz, N. 274
Porter, Douglas 12, 13, 15, 17, 95, 96, 97, 114, 120, 121, 136, 154, 156, 186, 187, 188, 189, 191, 194, 195, 196, 197, 199, 206, 236, 247, 425, 453, 523, 529, 533, 540, 541, 542, 546, 548, 556, 557, 561, 628, 630, 672, 673, 674, 675
Pressey, Sidney, L. 2, 3, 4, 6, 8, 9, 10, 22, 23, 24, 25, 29, 31, 32, 33, 34, 35, 42, 47, 59, 60, 66, 68, 69, 89, 94, 95, 115, 116, 120, 121, 124, 125, 127, 128, 129, 130, 138, 139, 215, 257, 260, 261, 299, 389, 393, 395, 399, 400, 418, 425, 453, 497, 501, 502, 522, 525, 532, 533, 540, 542, 543, 544, 548, 575, 586, 590, 605, 623, 624, 625, 628, 630, 644, 650, 653, 672, 674, 675, 679, 680, 688, 689
Price, G. R. 675

Q

Quisenberry, K. S. 647

R

Ramo, Simon, 3, 23, 27, 28, 29, 263, 312, 367, 390, 392, 540, 660, 663, 664, 675, 676, 685
Randall, S. 265, 269
Rath, G. J. 20, 288, 532, 546, 628, 676
Ratliff, Floyd 102
Reid, L. S. 222, 628, 677
Renshaw, S. 115
Richards, S. E. 182
Rigney, J. W. 2, 575, 594, 595, 597, 608, 609, 610, 625, 638, 677
Roberts, A. B. 118
Rogers, M. S. 452
Roshal, S. M. 259, 547, 649
Rothkopf, E. Z. 13, 258, 262, 305, 318, 319, 323, 397, 526, 529, 621, 628, 650, 677, 678
Russell, W. A. 420

S

Saltzman, I. J. 546, 678
Sapon, Stanley M. 96, 136, 154, 156, 173, 180, 182, 425, 470, 561, 592, 621
Saupe, J. L. 605, 610
Sawyer, M. O. 34, 644
Schaughency, M. D. 115
Schinnerer, Otto P. 179
Schlosberg, H. 123, 124, 125, 560
Schoellkopf, Judith 247
Schoenfeld, W. N. 221
Schuller, C. F. 114, 118
Schuster, D. H. 596
Schutz, R. E. 677, 678
Severin, D. G. 34, 675, 678
Sheffield, F. D. 17, 259, 533, 561, 581, 622, 639, 641, 651, 655, 665
Shephard, A. H. 307
Shulman, I. 274
Sidman, M. 320

Silberman, H. F. 397, 426, 452, 534, 543, 605, 606, 607, 677, 680, 681
Sims, Matilda 452
Skinner, B. F. 3, 4, 10, 11, 12, 13, 15, 16, 17, 20, 21, 23, 24, 25, 26, 29, 30, 94, 95, 96, 97, 98, 99, 100, 104, 116, 118 121, 122, 123, 124, 125, 130, 133, 137, 140, 141, 143, 145, 155, 159, 162, 167, 170, 173, 185, 187, 189, 190, 191, 195, 197, 198, 208, 209, 215, 216, 219, 221, 223, 237, 247, 257, 260, 261, 263, 299, 307, 308, 312, 389, 390, 393, 395, 399, 402, 416, 418, 419, 422, 425, 426, 427, 437, 438, 445, 446, 453, 464, 475, 477, 481, 482, 484, 486, 491, 493, 495, 506, 507, 508, 509, 513, 516, 521, 523, 526, 528, 529, 532, 533, 534, 535, 536, 540, 542, 543, 545, 548, 549, 552, 553, 554, 556, 557, 558, 559, 561, 564, 589, 590, 599, 600, 605, 610, 611, 612, 613, 614, 615, 620, 624, 625, 626, 628, 629, 630, 632, 634, 635, 638, 639, 640, 641, 643, 645, 646, 650, 653, 654, 656, 664, 667, 675, 676, 678, 681, 682, 683, 684, 685, 686, 689, 690
Skinner, Haylcon 274
Smith, D. E. P. 626, 686
Smith, H. P. 115
Speer, G. S. 131
Spence, K. W. 686, 687
Spitzer, M. E. 547, 562, 604
Stannard, A. F. B. 320
Staugas, L. 452
Stelter, C. 620, 649
Stephens, A. L. 34, 89, 502, 504, 675, 687
Stoddard, A. J. 387
Stolurow, L. M. 687
Studebaker, J. W. 274
Summers, S. A. 594, 638
Svenson, D. W. 677
Symonds, P. 57

T

Tate, T. R. 115
Taves, M. J. 118
Taylor, E. A. 115
TerLouw, Adrian 392
Thomas, W. 452
Thorndike, E. L. 67, 90, 91, 93, 94, 402, 689
Tinker, M. A. 420
Tolman, E. C. 67
Torreyson, H. C. 115
Toynbee, Arnold 385
Troyer, Maurice E. 34, 66, 120, 125, 127, 563, 644
Trump, J. 387
Tuber, R. J. 452
Twyford, L. C. 686

U

Underwood, B. J. 688, 689

V

Van Horn, C. 595, 677
Van Ormer, E. B. 116, 194, 259, 657, 666
Votaw, D. F. 126

W

Walker, E. S. 300, 585
Warren, A. B. 123
Washburn, J. N. 57
Weimer, Paul K. 262, 311, 689
Westfield, R. L. 262, 659
Weston, T. 269
White, Harvey 9, 646
Whitehead, A. N. 394
Whitlock, C. P. 115, 131
Wilson, M. P. 320
Wittich, W. A. 114, 116
Wood, B. D. 500
Wood, McKinnon 348
Woods, G. G. 66, 126
Woodworth, R. S. 52, 123, 124, 126, 307, 308, 324, 560, 629
Wulff, J. J. 17, 259, 604, 616, 617, 646, 647, 651
Wyckoff, L. B., Jr. 102, 222, 527, 532

Z

Zeaman, David 10, 13, 306, 526, 628, 650, 689, 690